2020

Neuropsychological Domains: Comparability in Construct Equivalence Across Test Batteries

Meghan Ashley Terzulli

NEUROPSYCHOLOGICAL DOMAINS: COMPARABILITY IN CONSTRUCT EQUIVALENCE ACROSS TEST BATTERIES

A dissertation submitted in partial fulfillment
of the requirements for the degree of

DOCTOR OF PSYCHOLOGY

to the faculty of

DEPARTMENT OF PSYCHOLOGY

of

ST. JOHN'S COLLEGE OF LIBERAL ARTS AND SCIENCES

at

ST. JOHN'S UNIVERSITY
New York
by

Meghan Ashley Terzulli

Date Submitted: ________________

Date Approved: ________________

____
Meghan Ashley Terzulli

____
Samuel O. Ortiz

ABSTRACT

NEUROPSYCHOLOGICAL DOMAINS: COMPARABILITY IN CONSTRUCT EQUIVALENCE ACROSS TEST BATTERIES

Meghan Ashley Terzulli

The National Center for Education Statistics (NCES, 2016) reported an increase in the number of non-native English-speaking students in U.S. public schools as well as a frequent coexistent correlation with low-SES and poverty, but not for all racial or ethnic minority groups. Because it is well known that SES and language difference play an important role in academic achievement, it is imperative that school psychologists attend to these variables when considering the validity of obtained test scores and their support for subsequent diagnostic conclusions, especially when current rates of ELLs in special education suggest that evaluations are not necessarily providing unbiased results (NCES 2013). This trend remains troublesome despite advances in psychometrics and test development based on theoretical models of intelligence (i.e., CHC, Luria). However, use of tests from varied theoretical camps provides an additional challenge, as not all batteries measure constructs in similar ways (i.e., construct equivalence).

As a result, this study evaluated the comparability of construct equivalence on neuropsychological measures across batteries and tests, the extent to which typical neuropsychological domains vary according to how much "language" is used in the measurement of each domain, and the equivalence of scores when domains are assessed in high SES monolingual and bilingual populations in a sample of 252 school-age individuals who underwent evaluations in a private clinic. Results indicated that there is variation in how domains are constructed on certain batteries, confirming that for some tests there is not construct equivalence; high SES bilinguals and monolinguals seem to perform just as well on language tests; and that linguistic demand impacts bilinguals' performance. Post-hoc analyses indicated that the presence of a diagnosis sometimes indicated poorer performance on domain tasks. Implications include the need to consider the impacts of language, disability, and SES when evaluating bilingual students, as well as test selection during evaluation planning. Further research is needed to address the differences in performance for high and low SES bilinguals and address the possible presence of a "bilingual advantage."

DEDICATION

Daniel,

Thank you for being my ever-shining light and for lifting me up throughout this degree program, which took longer to complete than either of us anticipated. This body of work could not have been finished without your loving sacrifices. Not only did you choose to support me from this dream's inception, but you held my hand when it became difficult, wiped my tears when anxiety struck, and found solutions when I thought all was lost. You (courageously) said "I do" when I was only halfway through my degree program and your efforts to help me see this dream become reality were only redoubled. You push me constantly to be better than I am, and I am truly grateful to be your wife.

I hope that this project's completion is a sign to all that what we set out to accomplish alone is always infinitely more fulfilling when it is accomplished together.

With all my love. Always.

ACKNOWLEDGEMENTS

Firstly, I would like to thank God for giving me the grace, persistence, and strength to see this project through, despite all obstacles. Mom and Dad, thank you for your support and encouragement along this journey to PsyD. To all my siblings, family, and friends, I could not have completed this great task without your love. Thank you, specifically to Dr. Cristina Balducci, my friend and colleague, for your time, compassion, and unrelenting assistance in completing my degree requirements. You walked with me through the hills and valleys of our doctoral program and stood by me in my steepest climbs over the last three years to this summit. I could not have made it here without you.

To my husband, Daniel, thank you. Your love and indefatigable support from the inception of my dream to be "Dr. Meg" permeates this project. From your support of my decision to enroll in this program to your hard work as my "research assistant" inputting data and double checking excel formulas, you have contributed largely to my ability to prepare my dissertation defense and achieve my dream. Your sacrifice of love does not go unnoticed and I will be grateful always. Thank you and I love you.

Thank you to my supervisors, specifically to Dr. Marsha Vasserman and Dr. Daniela Montalto, from the NYU Child Study Center at Hassenfeld Children's Hospital for providing me with the patient repository from which my data was collected. I am grateful to my Committee Members, Dr. Dawn Flanagan and Dr. Robin Wellington, for your insight and support of this research project. Finally, "thank you" does not seem sufficient to express my immense gratitude to my mentor, Dr. Samuel Ortiz. Without your guidance, patience, and kind encouragement over the last nearly four years, I would not have reached this mountaintop moment.

TABLE OF CONTENTS

LIST OF TABLES

Chapter I

Introduction

According to the National Center for Education Statistics (NCES, 2016), the number of students who are non-native English speaking (i.e., English Language Learner [ELL] or bilingual[1]) in United States public schools continues to increase. For example, in 2003-2004, ELLs comprised about 8.8% of the school-age population. Current estimates place the number in 2013-2014 at about 9.3%. The number of ELLs in U.S. public schools also varies greatly by state and ranges from 0.7% (West Virginia) to 22.7% (California; NCES, 2016). In New York State, the number of ELLs in public schools ranges approximately from 6.0 to 9.9% (NCES, 2016). These statistics suggest that school psychologists are likely to be encountering bilingual students more frequently both in assessment and treatment settings.

Apart from language difference, there is often a concomitant correlation with low-SES and poverty but not for all racial or ethnic minority groups. While many ELLs in public schools live in families who fall into the lower SES range, they are not distributed equally. For example, Hispanics comprise 77.7% of the ELL population but 28% of those who are in low-SES categories, as indicated by participation in free and reduced lunch programs (NCES, 2017). Other groups, such as Chinese and Japanese, tend to be disproportionally over-represented in high-SES categories (NCES, 2017). Because it is

[1] "Whereas the terms *English language learner* (ELL) and *bilingual* are used interchangeably in this document, and whereas bilingual often refers to an individual with proficiency in two languages, our use of the term bilingual is general and intended to refer to all individuals with any degree of experience in and exposure to a language other than English, including children who enter the U.S. school system (ELLs) and for whom English was not the native or heritage language. We recognize that an individual need not be bilingual to be an ELL, and conversely, an individual need not be an ELL to be bilingual."

NASP Position Statement: Bilingual Services (2015) well known that SES plays an important and significant role in academic achievement (e.g., correlation between SES and SAT scores), it is imperative that school psychologists attend to this variable every bit as much, if not more, than language difference when considering the validity of obtained test scores and their support for subsequent diagnostic conclusions.

A failure on the part of professionals to consider language and SES differences can have dramatic outcomes, many of which have not always been positive for ELLs. In fact, current rates of disproportional representation of ELLs in special education (NCES, 2013) suggest that evaluations are not necessarily providing unbiased results and that interpretations regarding the diagnostic meaning of test scores lean more often toward intrinsic problems than extrinsic factors (e.g., second language learning). This trend remains troublesome despite advances in psychometrics and test development as well as the convergence of developers on a common theoretical model of human cognitive abilities known as Cattell-Horn-Carroll (CHC; Carroll, 1993; Cattell, 1971; Horn & Noll, 1997) theory. The same can be said for the emerging field of school neuropsychology where adherence to the Lurian (Luria, 1966, 1973, 1980) model as the basis of evaluation continues to possess far less agreement regarding the nature and constitution of the various domains of interest (Ardila, 1992). School neuropsychologists face many of the same problems confronting school psychologists when it comes to conducting comprehensive evaluations because the assessments used most commonly for both purposes cannot be strictly determined to be equivalent in their domain measurement.

Although there has been some attention in addressing the potential threat to the validity of language and language-related domains in assessment, the vast majority of investigations have done so under the CHC (Carroll, 1993; Cattell, 1971; Horn & Noll, 1997) theoretical framework that best organizes psychological assessments. However, the Lurian theoretical framework delineates a clear overlap in its "blocks" with several narrow abilities described in CHC (Flanagan, Alfonso, Ortiz, & Dynda, 2010). For example, the measurement of fluid reasoning is distinct in CHC theory but is subsumed in both Block 2 (Simultaneous and Successive Cognition) and Block 3 (Planning and Metacognition). This overlap is precisely what causes issue in examination of construct equivalence, as different tests that purport to measure the same domain may, in fact, be diluted or measuring a completely different construct. An analysis of this kind has not been investigated in the field and adds significant value to understanding the best assessments to use in developing a comprehensive evaluation plan.

Chapter II

Literature Review

It is evident that bilingual and monolingual individuals differ on myriad levels, including cultural experiences, language development, and socio-economic status (SES), which impact their academic performance and classroom behavior (Flanagan & Ortiz, 2007; Flanagan et al., 2013; Ortiz, Melo, & Terzulli, 2017; Ortiz, 2008, 2011; Ortiz, Devine, & Ortiz, 2016; Thomas & Collier, 2002). These differences imply that bilinguals and monolinguals cannot be evaluated the same way. Much research has been conducted on English Language Learners (ELLs) who are non-disabled, of average ability, with moderate to high English proficiency, and tested in English (Bialystok, 2001a; Cormier, McGrew, & Ysseldyke, 2014; Cummins, 1984; Flanagan & Ortiz, 2007; Flanagan et al., 2013; Kranzler, Flores, & Coady, 2010; Ortiz et al., 2017; Ortiz et al., 2016; Ortiz, 2008, 2011; Ortiz, Flanagan, & Alfonso, 2015; Rhodes, Ochoa, & Ortiz, 2005; Sattler, 2001; Thomas & Collier, 2002). In general, this research has yielded two major findings: (1) Native English speakers perform better than English learners at the broad ability level on standardized, norm-references tests of intelligence and general cognitive ability and (2) English learners tend to perform significantly better on nonverbal type tests than they do on verbal tests (Flanagan et al., 2013; Ortiz et al., 2017; Ortiz et al., 2016; Ortiz, et al., 2015). As shown above, data strongly supports that bilinguals' test performance is influenced by the degree of expected language proficiency in English and their cultural experiences and knowledge. This third principle, then, can be included in order to best understand how to evaluate bilingual students: (3) Test performance of ELLs is moderated by the degree to which a given test relies on or requires age- or grade-expected English language development and the acquisition of incidental acculturative knowledge (Ortiz et al., 2017; Ortiz et al., 2016).

Given this third principle, it would be expected that bilinguals would perform better on tasks that require less English language and cultural knowledge. Thus, it appears that their test performance falls on a continuum of attenuation of performance (Flanagan et al., 2013; Ortiz et al., 2017; Ortiz et al., 2016; Ortiz, et al., 2015). Research shows that tests requiring lower levels of cultural and linguistic knowledge result in higher mean scores for bilinguals (Flanagan et al., 2013; Ortiz et al., 2017; Ortiz et al., 2016; Ortiz, et al., 2015). Likewise, tests that require higher levels of cultural and linguistic knowledge result in lower mean scores (Flanagan et al., 2013; Ortiz et al., 2017; Ortiz et al., 2016; Ortiz, et al., 2015). Thus, the "average" level of performance for a bilingual individual is not the same as that of a monolingual and must be taken into consideration when determining "disordered" performance.

Bilingual individuals' language development has been shown to be different than that of monolingual individuals, which has significant implications related to the development of academic skills, such as reading, writing, listening, and classroom behaviors (Bialystok, 2001a, 2001b; Bialystok & Barac, 2012; Bialystok, Craik, & Luk, 2012; Feifer & DeFina, 2000, 2002; Foy & Mann, 2014; Kroll & Bialystok, 2013; Morton, 2010; Sattler, 2001). Language acquisition is a developmental process that is subject to the maturation of the brain (Bialystok, 2001a; Feifer & DeFina, 2000, 2002; Kroll & Bialystok, 2013; Morton, 2010). Because reading and writing are symbolic aspects of language development (Feifer & DeFina, 2000, 2002), they are also subject to this maturation process. Thus, based on this data, it is clear that language development is a key factor, which is often affected by local resources, parent education, and SES (Ortiz et al., 2013). For example, the presence of increased access to resources, such as parents speaking more frequently to their children, providing support by means of tutors, access to bilingual books and music, etc., as often found in high socio-economic status (SES) households, presents an advantage to bilinguals of high SES in terms of their language development, as compared to bilinguals of low SES (Ortiz et al., 2013).

When looking at the neuropsychological domains using the Lurian model (1966, 1973, 1980), there is no real research that examines the comparability of construct equivalence among tests purporting to measure similar domains. In addition, there is a lack of insight as to whether or not the neuropsychological domains are affected by developmental language proficiency or acculturative knowledge acquisition or moderated by SES. Moreover, there is a dearth of research to show the extent to which language proficiency itself influences measurement of the various neuropsychological domains other than language. For example, is measurement of executive functions, including but not limited to attention, set shifting, planning, and organization, in bilingual populations equivalent to and valid as compared to the measurement of these domains in monolingual, English speaking, populations?

Neuropsychology and neuropsychological assessments are becoming increasingly popular, given their ability to tease out more information regarding learning disabilities and instructional needs. This makes it imperative that there be research to guide neuropsychological practice so that routine assessments do not lead to erroneous evaluation outcomes, such as overrepresentation in special education and poorer academic performance. Thus, more research is needed to identify to what extent, if any, language affects performance on neuropsychological tests and to what extent, if any, it affects the comparability of the neuropsychological domains in high SES monolingual and bilingual populations.

Research in neuropsychological evaluation of bilinguals has begun to examine differences in their performance on various tasks. One of the major findings illustrated that bilinguals performed better than monolinguals on the Stroop Color-Word Test and other response inhibition tasks (Bialystok, Barac, Blaye, & Poulin-Dubois, 2010; Carlson & Meltzoff, 2008; Costa, Hernandez, Costa-Faidella, & Sebastian-Galles, 2009; Foy & Mann, 2014). The higher performance was attributed to differences in executive functioning, such as increased control of inhibition due to frequent suppression of one language system (Bialystok et al., 2010; Bialystok & Viswanathan, 2009; Carlson & Meltzoff, 2008; Costa et al., 2009; Foy & Mann, 2014). However, bilinguals were also found to perform better on neuropsychological tasks that do not require inhibition, such as Trail Making (Bialystok et al., 2010; Carlson & Meltzoff, 2008; Costa et al., 2009; Foy & Mann, 2014). More current research seems to revolve around bilinguals' increased efficiency in monitoring functions. For example, bilinguals seem to perform better than monolinguals on monitoring tasks and are able to do so with less activation in brain areas involving monitoring (Costa et al., 2009; Kroll & Bialystok, 2013; Morales, Gomes-Ariza, & Balo, 2013; Paap & Sawi, 2014).

Other research has focused on the developmental effects of bilingualism. In a study by Kovacs and Mehler (2009), seven-month-old bilingual infants were able to switch their anticipatory gaze toward an attractive stimulus more quickly than monolinguals. In addition, preschool-aged bilingual children exhibited positive effects on task switching and inhibitory control on visual and auditory tasks (Martin-Rhee & Bialystok, 2008; Foy & Mann, 2014). Moreover, it appears that children who are bilingual and bi-literate outperform their monolingual peers in 12th grade by 20 percentile ranks (Thomas & Collier, 2002). This increased efficiency on tasks, influenced by bilingualism, occurs even in individuals who learn a second language later in life (Bialystok, 2001a; Bialystok et al., 2008; Bialystok, Craik, & Ruocco, 2006; Collier, 1995).

Other research seeks to examine the improved efficiency in the working memory, meta-linguistic awareness, increased comfort with language in general, and improved reasoning efficiency and problem-solving ability in bilinguals (Bialystok, 2001b; Bialystok & Barc, 2012; Carlson & Meltzoff, 2008; Cormier et al., 2014; Costa et al., 2009; Templeton, 2012). Bilinguals appear to use working memory more often and more efficiently as a function of constantly mentally translating (Bialystok, 1999, 2010, 2011; Bialystok & Martin, 2004; Carlson & Meltzoff, 2008; Costa et al., 2009). In terms of meta-linguistics, bilinguals seem to be able to use more aspects of language in the service of other cognitive functions (Bialystok, 2001b, Bialystok & Barac, 2012). Furthermore, bilinguals are more comfortable using language to fit their needs, often "playing" with it to create words to best express themselves (Bialystok, 1999, 2001b; Bialystok & Barac, 2012). Bilinguals are also bicultural, which seems to give them the ability to approach tasks in many ways instead of being limited by one perspective (Bialystok, 1999, 2001b, 2011). In general, there is research that investigates the effects of bilingualism on different abilities and that research is unified in their assertions that there are significant developmental implications.

ELL students can take at least five to seven years to gain a level of proficiency (cognitive academic language proficiency, or CALP) necessary to achieve at a comparative level to native English speakers (Cummins, 1984). In three to five years, however, ELLs develop basic interpersonal communication skills (BICS), which shows a basic use and understanding of English but lacks the depth and breadth of knowledge necessary for academic success (Cummins, 1984). Thus, ELLs underperform, but, for various reasons (e.g., broken speech, speaking with an accent, etc.), school faculty may suspect underlying deficits and refer these students for special education evaluations, which almost always involve the use of standardized tests. Furthermore, the likelihood that high SES bilingual students develop BICS and CALP is arguably more likely, given the presence of added resources unavailable to bilinguals with low SES. For example, low SES bilingual students often come from families where parents have less education, are required to work more often, and have less time to devote to engaging linguistically with their children in their native or second language (Ortiz et al., 2013).

Many of the characteristics of impaired readers, for example, are considered “normal” for typically developing bilinguals (Ortiz, Douglas, & Feifer, 2013; Feifer & DeFina, 2000). For instance, poor decoding skills in an impaired reader suggests difficulty with phonological processing; while, poor decoding in older bilinguals may be attributed to limited exposure to sounds in early childhood (Ortiz et al., 2013; Feifer & DeFina, 2000). In addition to difficulty with decoding, impaired readers can be characterized by: (1) weak vocabulary, due to inadequate exposure, (2) difficulty reading strategically, due to a problem with fluid reasoning, (3) poor spelling, because of difficulty with visual memory, (4) many opportunities to read outside of school, but not sufficient to improve reading skills, and (5) a tendency to avoid reading because it is effortful and difficult, leading to poor motivation and low confidence (Ortiz et al., 2013; Feifer & DeFina, 2000). Many of these are characteristic of typically developing bilingual students but are the result of altogether different problems. Their weak vocabulary, for example, may be due to lack of English exposure, although their spelling would not suffer (Ortiz et al., 2013; Feifer & DeFina, 2000). Their inability to read strategically may be due to limited educational opportunity or benefit and insufficient reading opportunities (Ortiz et al., 2013; Feifer & DeFina, 2000). The outcome for these students, however, is the same: a tendency to avoid reading because it is effortful and difficult, leading to poor motivation and low confidence (Ortiz et al., 2013; Feifer & DeFina, 2000).

When reading for comprehension, bilingual students struggle to infer meaning because they lack the cultural knowledge and experience with the English language (Feifer & DeFina, 2000; Ortiz, 2016). More experience garners clearer meaning and better overall comprehension of the text (Feifer & DeFina, 2000; Ortiz, 2016). Moreover, monolingual English speakers typically cease "decoding" as they become more advanced readers and begin to recognize words based on their orthographic processing of letters, words, and sentences in order to derive meaning (Feifer & DeFina, 2000; Ortiz, 2016). Bilingual students have less experience and less ability to extract meaning using orthography automatically or fluently (Feifer & DeFina, 2000; Ortiz, 2016). In addition, students learning a second language hear and interpret the sounds they hear in a manner that conforms to words they already know (Feifer & DeFina, 2000; Ortiz, 2016). In this way, the brain attempts to make sense of what it is hearing and provide meaning (Feifer

& DeFina, 2000; Ortiz, 2016). Often, bilingual students' difficulties in these academic areas are labeled as disordered; whereas, in actuality, their struggle is typical for a developing bilingual (Flanagan et al., 2013; Ortiz, 2016; Ortiz, Flanagan, & Dynda, 2008).

Bilingual students, in addition to their academic needs, often present with what appear to be behavioral difficulties. Bilingual students may be slow to begin tasks because they have limited understanding of the classroom language or slow to finish tasks because of constant translation (Bialystok, 1999, 2011; Ortiz, 2016; Sattler, 2001). Because of their difficulty understanding the language of the classroom, bilingual students may not understand the classroom rules or norms, may have difficulty encoding information into memory, and may attempt to discuss with other students to attempt to understand tasks or instructions (Flanagan, Alfonso, Ortiz, & Dynda, 2008, 2010; Ortiz, 2016; Rhodes et al., 2005; Sattler, 2001). These difficulties and attempts at comprehension may be seen by the classroom teacher as forgetfulness, inattention, distractibility, impulsivity, or disruptiveness (Flanagan, Ortiz, Alfonso, & Mascolo, 2006; Ortiz, 2016; Rhodes et al., 2005; Sattler, 2001). However, many of the issues bilingual students face behaviorally can be easily addressed with academic supports.

To summarize, psychologists in the United States judge the performance of bilinguals using standards that were created to evaluate monolingual and monocultural individuals (Bialystok, 2001a; Cormier et al., 2014; Cummins, 1984; Flanagan & Ortiz, 2007; Flanagan et al., 2013; Kranzler et al., 2010; Ortiz et al., 2017; Ortiz et al., 2016; Ortiz, 2008, 2011, 2016; Ortiz et al., 2015; Rhodes et al., 2005; Sattler, 2001; Thomas & Collier, 2002). This is extremely problematic because, as is evident, bilingual and bicultural students have vastly different experiences than monolingual and mono-cultural students, which provide implications for their learning and classroom needs. When evaluating bilingual students, it is imperative to take into consideration the developmental language proficiency and acculturation (Flanagan et al., 2013; Ortiz et al., 2017; Ortiz et al., 2016; Ortiz, 2016; Ortiz et al., 2015). Both are developmental processes and affect age-based expectations of performance (Bialystok, 2001a; Feifer & DeFina, 2000, 2002; Flanagan et al., 2013; Ortiz et al., 2017; Ortiz et al., 2016; Ortiz, 2016; Ortiz et al., 2015). In addition, taking a student's SES into account is equally important in highlighting language development patterns and needs. Overall, early language development has long-lasting effects that manifest in evaluations with bilingual students and create cognitive and behavioral differences that imitate disorders.

For the reasons stated, the validity of standardized test batteries in the assessment of bilinguals has been called into question by numerous researchers over the decades (Comier et al., 2014; Cummins, 1984; Flanagan et al., 2008; Ortiz, 2008, 2011; Valdes & Figueroa, 1994). Standardized tests pose a validity problem for ELLs because the obtained results are at risk for representing the extent to which culture and language had on their performance, as opposed to the constructs the tests were intended to measure (Comier et al., 2014; Cummins, 1984; Flanagan et al., 2008; Kranzler et al., 2010; Ortiz, 2008, 2011). Thus, the obtained test scores likely under-represent their actual abilities, yet, these scores are still used to make special education decisions, leading to the over-classification and representation of ELLs in special education (Comier et al., 2014; Cummins, 1984; Flanagan et al., 2008; NCES, 2013; Ortiz, 2008, 2011). It is imperative, then, to identify the "typical" performance of ELLs in order to understand how they perform on standardized tests and allow for the systematic evaluation of validity in current evaluations (Flanagan et al., 2013; Ortiz, 2008, 2011; Ortiz et al., 2015).

Nondiscriminatory assessment practices have been developed through the lens of CHC theory in order to evaluate ELLs in a way that will yield the most valid results (Ortiz, 2008, 2011; Ortiz et al., 2008; Ortiz et al., 2013). Ortiz (2008) lays out a framework for evaluators to use when assessing children who are experientially and/or linguistically different than those brought up in mainstream American culture. His framework proposes that evaluators go through a number of evaluative steps before considering a formal evaluation with standardized testing. In general, these steps include assessment of alternative measures, such as curriculum based measures or work samples; learning ecology, such as the goodness of fit between the student and teacher; and language proficiency in the native language and in English; selection of tests that are most appropriate to the referral concern and that focus on assessing the specific constructs in question, in addition to those that provide broad general information about functioning, interpretation of results within the context of the individual's unique educational, experiential, and familial background, and conclusions based on multiple sources of information.

The procedures described in this framework, however, are effective for general psychological and psychoeducational evaluations conducted under the CHC framework, but are not when conducting neuropsychological evaluations, as the domains and constructs are different, even combined. In the CHC framework, the domains can be considered distinct from one another, such as visual-spatial abilities and fluid reasoning abilities (Carroll, 1993; Cattell, 1971; Horn & Noll, 1997; Flanagan & McGrew, 1998;

Flanagan & Ortiz, 2007; Flanagan, et al., 2000; Flanagan et al., 2013; Flanagan et al., 2006; McGrew & Flanagan, 1998). The neuropsychological domains and constructs, however, incorporate a variety of different abilities (Ardila, 1992; Flanagan et al., 2010; Luria, 1966, 1973, 1980; Miller, 2013, 2015). For example, nine CHC domains are subsumed into one Luria Block, including visual-spatial, fluid and quantitative reasoning, auditory processing, short-term and long-term memory, crystallized and educational knowledge, and processing speed (Flanagan et al., 2010; Luria, 1966, 1973, 1980; Miller, 2013, 2015; Ortiz, 2016). This overlap is precisely the reason for the need to investigate the batteries and tests being used in evaluations, as they may be intending to measure a given construct, but different tests may measure the same construct in non-equivalent ways.

Within the CHC framework, subtests from the major cognitive batteries are classified according to their degree of cultural and linguistic loading; this classification system became known as the Culture-Language Test Classifications (C-LTC; Flanagan, McGrew, & Ortiz, 2000; McGrew & Flanagan, 1998). Using these classifications, Flanagan and Ortiz (2001), Flanagan, Ortiz, and Alfonso (2007), and Flanagan, Ortiz, and Alfonso (2013) developed the Culture-Language Interpretive Matrix (C-LIM), which was designed to serve as a practical tool for clinicians “to evaluate the extent to which differences in developmental language proficiency and acculturative learning opportunity may have affected the validity of scores obtained from standardized tests” (Flanagan, Ortiz, & Alfonso, 2013, p 309). The C-LIM is based on the three previously described principles, as outlined by Ortiz, Flanagan, and Alfonso (in press). Using these principles, the developers created a matrix that consists of nine cells representing varied degrees (low, moderate, and high) and combinations of cultural and linguistic loading. Subtests in the top left corner of the matrix are expected to have the lowest degree of cultural and linguistic loading, where ELLs are expected to perform at or near the mean. On the contrary, subtests in the bottom right corner of the matrix are expected to have the highest degree of cultural and linguistic loading, and ELLs are expected to perform most poorly on these subtests. The horizontal axis represents increasing linguistic demand, while the vertical axis represents an increase in cultural demand. Diagonal downward movement from the top left to the bottom right corner represents the combined effect of cultural and linguistic loading on test scores.

Using the available research on ELL test performance, the C-LIM generates an expected pattern of attenuated performance for ELLs based on their degree of cultural and linguistic difference, which appears in the matrix as a systematic declining pattern. If this systematic declining pattern of performance is consistent with other sources of information gathered, then the obtained test scores are deemed invalid and uninterpretable, indicating that the pattern of decline is determined to be *primarily* due to the effects of culture and language rather than extrinsic (i.e., environmental, behavioral) or intrinsic (i.e., emotional, disability) factors. However, if the pattern is inconsistent with the expected pattern of decline, then it can reasonably be deduced that some other factors are likely accounting for the ELLs' performance, and culture and language are only *contributory* factors. Of particular note is that based on research (i.e., Sotelo-Dynega et al., 2014), the expected pattern of performance in the C-LIM may be adjusted according to factors that may render the examinee just "slightly" different from monolingual peers or "moderately" different, or even "markedly" different. Whereas the predominant factor in making such a determination lies with the amount of developmental exposure to English, it also alludes to the presence of SES, particularly as it may influence education and development of one's own heritage language even in the absence of formal bilingual education. This suggests that the higher the SES of an examinee, the less the expected effect on language differences among bilinguals and monolinguals, particularly on language-based tasks. This can also be a consideration in the so-called "bilingual advantage."

This kind of systematic paradigm is progressive when looking at ELL performance on cognitive, CHC-based, tasks; however, because school-based neuropsychological evaluations do not adhere to the CHC framework for evaluation or interpretation of test results, there is no mechanism for evaluating the impact of cultural and linguistic variables on the measurement of the typical Lurian blocks which are, by CHC terms, "messy" and intentionally overlapping in the abilities that comprise them (Carroll, 1993; Cattell, 1971; Horn & Noll, 1997; Flanagan et al., 2010; Luria, 1966, 1973, 1980; Miller, 2013, 2015; Ortiz, 2016). Thus, to what extent such extraneous factors, such as developmental language proficiency, differentially affect the manner and comparability in which such neuropsychological abilities are constricted and measured in both monolinguals and bilinguals remains unanswered.

Chapter III

Research Questions and Hypotheses

Based on the preceding discussion of research on neuropsychological evaluation of monolingual and bilingual individuals, it seems clear that there is variable consistency of the measurement of various domains because neuropsychologists typically use many tests to measure the same domains. Specific questions to be addressed by the study include:

1. What is the current degree of comparability of neuropsychological measures across different batteries and tests?
2. To what extent do the typical neuropsychological domains vary according to how much "language" is used in the measurement of each domain?
3. When the same neuropsychological domains are assessed in high SES monolingual and bilingual populations, do the scores remain equivalent or are they potentially affected by variation in characteristics unique to the specific combination of tests being used (i.e., language)?

Hypotheses

Based on the nature of the research questions that were elicited from the literature on the topic of neuropsychological evaluations of monolinguals and bilinguals, as impacted by SES, the following hypotheses are presented as testable propositions to which the current study will address itself. These hypotheses include:

1. The null assumption is that different neuropsychological tests represent and construct the same neurocognitive domains (i.e., memory, attention, fluid reasoning, etc.) in similar ways on similar (i.e., cognitive, neuropsychological,

etc.) tests. Alternatively, it is possible that there is variation in how the neurocognitive domains are represented and constructed in neuropsychological tests and batteries as compared to other similar tests and batteries.

2. The null assumption is that, given the prevailing research, there is substantial variation within the neuropsychological domains according to how much "language" is used in the measurement of each one, the measurement of language notwithstanding, in both bilinguals and monolinguals of high SES backgrounds. Alternatively, it is possible that, unless language is the domain being measured, language is an irrelevant influence in the measurement of the other neuropsychological domains, among high SES bilinguals and monolinguals.
3. The null assumption is that when the same neuropsychological domains are assessed in monolingual and bilingual populations, the scores of assessed monolinguals and bilinguals will be greatly influenced by the variations in language loadings within each neuropsychological domain.
   Alternatively, it is possible that, even when assessing domains other than language, the scores remain equivalent and are not influenced significantly by variations in language, other than in the assessment of the language domain itself, given the high SES backgrounds of the monolingual and bilingual individuals.

Chapter IV

Methods

Participants

Participants were children and adolescents aged 3 years to 21 years old who lived in or near New York City and who underwent neuropsychological evaluations at the Child Study Center at Hassenfeld Children's Hospital at New York University Langone Medical Center. Private evaluations at this clinic are costly and not subject to insurance and are, thus, typically paid for by families out of pocket.

Inclusionary and Exclusionary Criteria. In order to participate in this study, the participants were ELLs, who self-reported knowledge or experience with a language other than English, or monolinguals, who self-reported knowledge or experience with the English language only. Participant data must have been available for all aspects of the evaluation and participant reports available for review. Participants were excluded from this study if they did not speak any English or if their evaluation did not include one or more of the following assessments: intellectual, achievement, executive functioning, or behavioral. Participants were also excluded from the study if their adaptive skills were in the below average range based on scores reported on the Adaptive Behavior Assessment Scale, 2nd Edition (ABAS-II; GAC of 70 or below) and/or Vineland Adaptive Behavior Scales, 2nd Edition (Vineland-II; ABC of 70 or below) and/or Behavior Assessment System for Children, 3rd Edition (BASC-3, Adaptive Composite of 30 or below).

Group Assignments. Participants were randomly selected from the database and intentionally assigned to a monolingual and bilingual group, based on their self-reported language status. Participants were matched by age, gender, and disability diagnosis.

Settings

All evaluations were completed at the Child Study Center at Hassenfeld Children's Hospital at NYU Langone Medical Center prior to December 2017. The data was housed in a secure and locked computerized patient registry within this facility, in the writer's onsite supervisor's office. All data collection was completed onsite. The deindentified collected data was transferred to a separate, secure and locked, computerized database in this writer's possession.

Consent

The patient registry in question is currently under an Institutional Review Board (IRB) associated with the Child Study Center. According to the criteria set for the data in the patient registry, the data collected from neuropsychological evaluations at the facility and included in the registry do not require consent, informed or otherwise. Therefore, consent would not be required retroactively for those participants' data selected for inclusion in the study. Individuals who are evaluated at this facility, however, are required to obtain parental consent to be evaluated using a facility-approved consent form. Children's verbal assent is also obtained to participate in the evaluation process. IRB approval was obtained from NYU Medical Center to conduct the retroactive chart review in order to determine eligibility of all prospective participants. Overall, minimal risk is involved in participation in this study.

Data Collection and Procedures

Data was collected and analyzed retrospectively from the patient registry. The writer screened participants' evaluation reports and demographics forms to code for self-reported bilingual status by obtaining their names from the patient registry, accessing their electronic patient charts, and reviewing the evaluation reports and forms. Once bilingual status was coded, all the data was de-identified, including no personal identifying information, and a new database was created. The data was then screened for inclusion in the study and the data of all participants who met the inclusionary criteria was included in the final study database. All participants' data remained de-identified and were given an identification number for the study.

Demographic data, including but not limited to SES, age, grade, race/ethnicity, sex, handedness, diagnoses, speech or language delays, and motor delays were included in the original patient registry. As such, demographic data was not actively collected from participants' records by the writer but was included in the study's analyses. Assessment data was also included in the original patient registry and was collected for included participants only by the writer into the study database, as described above.

Measures

Measures given throughout the course of the previously given evaluations may have include, but are not limited to, standardized test batteries, rating scales, structured/semi-structured/unstructured observations, and clinical interviews. These varied for each participant, depending on age and referral concern. It is important to note that many of the tests overlap in the domains they are designed to measure, in addition to the CHC domains (see Figure 1 above). Specific measures that were coded and evaluated include the following: Intelligence Tests, Achievement Tests, Tests of Attention and Executive Functioning, Learning and Memory Tests, Language Tests, Visual-Spatial Tests, Motor Tests, and Adaptive Measures.

Chapter V

Results

After data collection, all of the data were entered into the Statistical Package for the Social Sciences (SPSS 26.0), where descriptive statistics, paired-samples *t* tests, and independent samples *t* tests ($p < .05$) were conducted.

This results section is divided into four sections. The first section describes the preliminary data analysis which took place prior to the primary analyses of hypothesis testing. The second section described the characteristics of the participants that were included in the investigation. The third section delineates analyses related to the three hypotheses posed. Namely, that (1) there is variation in how the neurocognitive domains are represented and constructed in neuropsychological tests and batteries as compared to other similar tests and batteries; (2) that measurement of the language domain likely leads to greater variation in performance within the neuropsychological tests beyond actual ability and as a function of how much "language" is used in its measurement of each one; and (3) that when assessing domains other than language, the scores derived on monolinguals and bilinguals will be affected by the relative differences in the degree to which language is used in the measurement of these non-language domains. Finally, the fourth section describes post-hoc analyses that were conducted following a review of the results of primary analyses and the development of additional questions to be addressed.

Preliminary Data Analysis

Preliminary data analysis was conducted using several statistical analyses. Descriptive and frequency analyses were run for all demographic data. Due to the lack of research to guide the selection of sample size, a target range of 100 participants in each group, for a total of 200 participants was selected by the writer and her faculty advisor. Initial data comprised a total of 267 participants based upon inclusion criteria during data collection stages, including 62 bilingual participants and 205 monolingual participants. Following initial analyses and consultation with this writer's faculty advisor, several participants were excluded from further study analyses, including a participant aged 22 years (ID 123), who was outside the ages to be included in the study, and 13 participants who had varying levels (i.e., mild to profound) of hearing impairment (ID 1, 124, 130, 131, 134, 135, 143, 146, 155, 179, 239, 240, 261), as their assessments scores would likely skew results in the language domains being evaluated.

Participant Characteristics

Final data to be included in study analyses comprised a total of 252 participants, 61 bilingual (24.2%) and 191 monolingual (75.8%) subjects. Demographic and frequency analyses were conducted for all included participants. Participants included in the investigation ranged in age from 3 years to 21 years of age, with a mean age of 10.41 years. Participants were found to range in grade level from Pre-Kindergarten to College, with a mean grade of 5.04. Most participants were Caucasian (72.5%), followed by Latino (8.6%), African American (7.0%), Mixed (6.6%), Asian (3.7%), and Other (1.6%). Further, participants were 58.3% male and 41.7% female. Participants also had a variety of diagnoses, including neurocognitive disorders, speech or language disorders, and motor disorders. Table 1 includes a list of the disorders and the frequency of their presence in the sample population.

Participants were noted to speak varied languages and emanate from numerous countries. Bilingual participants (24.2%) reported speaking at least one language other than English. Seventeen languages were represented with Spanish being the most frequently reported (11.5%). The remaining languages included: Mandarin (1.6%), Hebrew (1.6%), Italian (1.6%), French (1.2%), Russian (1.2%), German (0.8%), Greek (0.8%), Arabic (0.8%), Swedish (0.4%), Punjabi (0.4%), Polish (0.4%), Tagalog (0.4%), Gujarati (0.4%), Czechoslovakian (0.4%), Ethiopian (0.4%), and Portuguese (0.4%). In addition, four participants reported speaking a third language, including Mandarin (2 participants), German (1 participant), and Swedish (1 participant). In addition, 13 participants reported emigrating from a different country, including: Germany, Israel, Guatemala, China, Ecuador, Czech Republic, Dubai (UAE), Canada, England (UK), Ethiopia, and Brazil. Of these 13 participants, years of residence in the United States ranged from 6 to 16 years with a mean of 9.92 years. Furthermore, years of residence in their native country prior to immigration to the United States ranged from 0 to 7 years with a mean of 2 years.

Table 1

Participant Characteristics: Frequency of Presenting Diagnoses

Diagnosis	Frequency	Percent
Major Neurocognitive Disorder Without Behavioral Disturbance	2	.8
Other Persistent Mental Disorder Due to Conditions Classified Elsewhere	1	.4
Cognitive Disorder NOS	8	3.2
Major Depressive Affective Disorder, Single Episode, Unspecified	1	.4
Major Depressive Affective Disorder, Single Episode, Moderate	1	.4
Major Depressive Affective Disorder, Single Episode, In Partial Remission	1	.4
Major Depressive Affective Disorder, Recurrent Episodes, Moderate	1	.4
Major Depressive Affective Disorder, Recurrent Episodes, In Partial Remission	3	1.2

Diagnosis	Frequency	Percent
Unspecified Episodic Mood Disorder	6	2.4
Autistic Disorder, Active State	10	4.0
Other Specified Pervasive Developmental Disorder, Active State	3	1.2
Anxiety, Unspecified	24	9.5
Panic Disorder Without Agoraphobia	1	.4
Generalized Anxiety Disorder	15	6.0
Other Specified Anxiety Disorder	3	1.2
Social Phobia	6	2.4
Obsessive Compulsive Disorder	3	1.2
Gender Identity Disorder of Children	1	.4
Tic Disorder, Unspecified	2	.8
Transient Tic Disorder	2	.8
Chronic Motor or Vocal Tic Disorder	2	.8
Tourette's Syndrome	3	1.2
Encopresis	1	.4
Other and Unspecified Special Symptoms or Syndromes Not Elsewhere Classified	9	3.6
Adjustment Disorder with Depressive Mood	2	.8
Separation Anxiety	2	.8
Adjustment Disorder with Mixed Anxiety and Depressed Mood	1	.4
Depressive Disorder Not Elsewhere Classified	15	6.0
Oppositional Defiant Disorder	10	4.0
ADHD-Inattentive Type	75	29.8
ADHD-Hyperactive/Impulsive Type	81	32.1
ADHD NOS	20	7.9
Specific Learning Disorder (SLD) Reading	60	23.8
Specific Learning Disorder (SLD) Math	36	14.3
Specific Learning Disorder (SLD) Writing	35	13.9
Development Speech or Language Disorder	1	.4
Expressive Language Disorder	3	1.2
Mixed Receptive-Expressive Language Disorder	13	5.2
Other Developmental Speech or Language Disorder	16	6.3
Developmental Coordination Disorder	91	36.1
Other Specified Delays in Development	3	1.2
Unspecified Delays in Development	41	16.3
Mild Intellectual Disability	2	.8
Mild Neurocognitive Disorder	1	.4

Analyses

Hypothesis 1. The first hypothesis posits that there is variation in how the neurocognitive domains are represented and constructed in neuropsychological tests and batteries as compared to other similar tests and batteries. The null hypothesis is that there is no variation in domains regardless of how they are comprised. In order to test this hypothesis, data (i.e., subtest and/or index scores) were grouped by battery into the neuropsychological domains they purported to measure. These scores were used to calculate a mean which would define each battery. Finally, a paired-samples *t* test was conducted to compare the mean scores from one battery to another, the results of which can be found in Table 2.

Significant differences were found in the following domains: Fine Motor, Executive Functions on the DKEFS as compared to the Stroop Test, Cognitive Efficiency on Trails as compared to Digit Span and on Digit Span when compares to Auditory Consonant Trigrams (ACT). These suggest that there is variation in how the domains are constructed on certain batteries. Note, pairs with less than 25 cases were not included in the results due to their lack of robustness. In addition, there were no valid pairs to complete analyses on the following domains, and they are, thus, not reported: Expressive Language, Receptive Language, Learning Efficiency, Retrieval Fluency, Speed of Lexical Access.

Table 2

Comparison of Battery Means to Determine Comparability of Measurement

	Mean Diff.	*t*	*df*	Sig. (2-tailed)
Pair 1 (Visual Motor Integration) Beery – NEPSY	1.482	1.304	63	.197
Pair 2 (Fine Motor) Grooved Peg. – Perdue Peg.	-11.423	-6.905	249	.000*
Pair 3 (Fine Motor) Perdue Peg. – NEPSY	10.368	8.703	110	.000*
Pair 4 (Fine Motor) Grooved Peg. – NEPSY	1.270	.511	110	.000*

	Mean Diff.	*t*	*df*	Sig. (2-tailed)
Pair 5 (Retrieval Fluency) STROOP – DKEFS	-11.497	-6.398	66	.000*
Pair 6 (Speed of Lexical Access) DKEFS – STROOP	11.669	6.513	66	.000*
Pair 7 (Delayed Verbal Memory) CVLT – WRAML	-.774	-.473	88	.637
Pair 8 (Immediate Visual Memory) RCFT – WRAML	1.209	.414	51	.681
Pair 9 (Delayed Visual Memory) RCFT – WRAML	1.204	.604	47	.549
Pair 10 (Executive Functions) DKEFS – STROOP	11.328	4.829	50	.000*
Pair 11 (Executive Functions) STROOP – TOL2	-2.841	-1.630	76	.107
Pair 12 (Executive Functions) TOL2 – BIBER	-1.816	-1.748	178	.082
Pair 13 (Executive Functions) BIBER – NEPSY	1.471	.964	106	.337
Pair 14 (Executive Functions) WCST – DKEFS	-6.648	-1.783	27	.086
Pair 15 (Working Memory) WISC – DIGIT SPAN	1.376	1.466	123	.145
Pair 16 (Cognitive Efficiency) TRAILS – DIGIT SPAN	-4.388	-4.093	123	.000*
Pair 17 (Cognitive Efficiency) DIGIT SPAN – ACT	14.782	6.051	46	.000*

An asterisk (*) denotes statistical significance at the $p < .05$ level.

Hypothesis 2. The second hypothesis suggests that measurement of the language domain likely leads to greater variation in performance within the neuropsychological tests beyond actual ability and as a function of how much “language” is used in its measurement of each one. The null hypothesis is that only true ability, not language, impacts performance in the language domain. In order to test this hypothesis, the sample was divided into bilingual and monolingual groups and an independent-samples *t* test was conducted on the following language domains only: Expressive Language and Receptive Language, using the mean scores previously calculated for each battery.

Significant differences were not found in this analysis. In fact, Expressive

Language as measured on the Clinical Evaluation of Language Fundamentals (CELF), 5^{th} Edition (p = .087) suggested that bilinguals (M = 98.69) outperform monolinguals (M = 88.50). This contradicts the prevailing literature and led to the development of further questions and conduction of additional analyses, which will be addressed below (see Post Hoc Analyses). In addition, it supports the assumption that high SES bilinguals perform just as well, if not better than, high SES monolinguals.

Hypothesis 3. The third hypothesis states that when assessing domains other than language, the scores derived on monolinguals and bilinguals will be affected by the relative differences in the degree to which language is used in the measurement of these non-language domains. The null hypothesis is that only true ability, not language, affects performance in non-language-based domains. To test this hypothesis, an independent-samples *t* test was conducted on the remaining domains using the previously calculated mean scores for each battery, the results of which can be found in Table 3.

Significant differences were found in the following domains: Delayed Visual Memory as measured on the Rey Complex Figure-Drawing Test (RCFT), Executive Functions as measured by the Tower of London Test, 2^{nd} Edition (TOL-2), and Cognitive Efficiency as measured by Auditory Consonant Trigrams (ACT). This suggests that the amount of language included (i.e., receptive or expressive language) in the subtests measuring these specific domains on the indicated batteries impacts performance.

Table 3

Impact of Language on Subtest Performance by Domain

	Bilingual		Monolingual		*t*	*p*
	M	*SD*	*M*	*SD*		
Visual-Motor Integration: Beery	93.96	8.335	93.18	10.505	.451	.652
Visual-Motor Integration: NEPSY	96.59	13.535	91.62	12.074	1.578	.119
Fine Motor: Grooved Pegboard	82.36	33.388	89.22	22.518	-1.82	.070
Fine Motor: Perdue Pegboard	99.24	4.185	98.88	5.893	.432	.666
Fine Motor: NEPSY	87.93	10.816	88.38	11.556	-.184	.854
Learning Efficiency: CVLT	100.74	10.815	97.07	11.221	.173	.084
Learning Efficiency: WRAML	99.00	13.816	102.63	11.100	-.95	.346
Retrieval Fluency: NEPSY	97.50	10.607	102.30	12.182	-.998	.326
Retrieval Fluency: STROOP	96.28	10.368	93.90	11.664	.878	.382
Retrieval Fluency: DKEFS	99.39	14.413	100.85	13.095	-.612	.514
Speed of Lexical Access: DKEFS	103.40	12.442	102.72	12.050	.317	.752
Speed of Lex. Access: STROOP	96.28	10.368	93.90	11.664	.878	.832
Speed of Lexical Access: NEPSY	97.50	10.607	102.30	12.182	-.998	.326
Imm. Verbal Memory: CVLT	100.10	11.127	97.44	13.059	1.109	.269
Imm. Verbal Memory: WRAML	99.69	13.628	104.70	11.606	-1.38	.175
Delayed Verbal Memory: CVLT	100.80	12.645	98.75	12.594	1.105	.270
Delay. Verbal Memory: WRAML	97.16	16.662	99.92	14.235	-.773	.442
Immediate Visual Memory: RCFT	82.25	21.001	87.49	17.938	-1.90	.058
Imm. Visual Memory: WRAML	94.67	8.121	87.84	12.979	1.887	.065
Delayed Visual Memory: RCFT	88.27	13.413	93.11	12.906	-2.51	.013*
Delay. Visual Memory: WRAML	97.33	14.407	93.81	9.043	1.048	.300
Executive Functions: DKEFS	100.87	14.051	100.94	11.175	-.031	.097
Executive Functions: STROOP	93.67	8.315	92.43	11.051	.501	.618
Executive Functions: TOL-2	89.40	9.801	94.50	11.733	-2.69	.008*
Executive Functions: BIBER	95.85	8.489	96.36	7.959	-.427	.670
Executive Functions: NEPSY	96.99	14.345	93.93	14.563	.923	.358
Executive Functions: WCST	89.6	12.951	97.87	16.302	-1.40	.174
Working Memory: WISC	95.88	15.809	100.33	15.718	-1.26	.209
Working Memory: DIGIT SPAN	97.27	11.589	97.88	15.739	.094	.875
Cognitive Efficiency: DKEFS	94.86	15.945	96.00	15.843	-.243	.810
Cognitive Efficiency: NEPSY	95.83	14.634	92.41	15.964	.483	.633
Cognitive Efficiency: TRAILS	95.54	10.016	95.75	16.629	-.053	.958
Cog. Efficiency: DIGIT SPAN	97.22	10.498	100.45	13.291	-.978	.330
Cognitive Efficiency: ACT	94.05	11.402	86.28	14.779	2.092	.040*

An asterisk (*) denotes statistical significance at the p < .05 level

Post-Hoc Analyses

As mentioned above, several additional questions and hypotheses arose upon review of this study's primary analyses. Specifically, there were three additional questions to be addressed:

1. Is there any variability in the created composite scores that might explain the narrow range of scores seen, specifically on the language constructs of expressive and receptive language, in the results of primary analyses?
2. There appear to be subjects that have scores which indicate the presence of a speech or language disorder but are not counted in the analyses as having a disorder. Similarly, there appear to be subjects that have average or above average language scores and are counted as having a speech or language disability. As such, in order to more accurately portray language difficulties, subjects' disability status should concur with their language scores. Therefore, can adjustments be made to better represent language difficulties as a whole in follow-up analyses?
3. Does homogenization of the groups by diagnosis classification (i.e., Anxiety, ADHD, SLD, Speech/Language) aid in understanding the impact of language on subject performance?

In order to investigate the question of composite score cohesion, an analysis of subtest scores' variability was conducted to determine whether subtest scores which comprised the expressive and receptive language composites, by battery, demonstrated variability of more than one standard deviation (Standard Score Mean = 100, SD = 15). The analysis revealed that all subtests included in composite scores for receptive and expressive language domains for each battery evaluated were cohesive. In other words, there was no significant variability in subtest scores that might lead us to believe that the composites were equalizing polarized performance. This supported the hypothesis that high functioning, high SES, bilinguals, impact the narrow range of scores seen in the data collected, which is contrary to the expectation of bilingual versus monolingual performance in general, based upon the prevailing research.

In addressing the second point, where the diagnosis of a speech or language disorder was inconsistent with the subject's test performance, adjustments were made to identify subjects with language scores below a standard score of 90 as having a speech or language disorder and those with scores above 90 as not having a disability. When these adjustments were made, seven total subjects (three bilingual) had below average (< 85) scores that did not have a diagnosis of a speech or language impairment. These were re-coded to better reflect language abilities in analyses. After this adjustment, there are a total of 33 monolingual participants and 7 bilingual participants with a speech or language diagnosis included in the sample. Therefore, although the intent of this study was to address and examine the impact of language on the measurement of each construct, the available sample of subjects was not sufficient to accomplish this goal. Thus, the focus of this study has shifted to examine the impact of a given a diagnosis on neuropsychological test performance in a high SES bilingual population.

In order to homogenize the groups, independent samples *t* tests were conducted for each domain constructs and grouped by diagnosis category (i.e., Anxiety, ADHD, etc.). Specific diagnoses were selected based upon the number of subjects pertaining to the group in order to have a large enough sample size to evaluate. These diagnoses include, Anxiety, ADHD (which includes all subtypes: Inattentive, Hyperactive/

Impulsive, and Combined), Specific Learning Disorder (SLD): Reading, SLD: Math, SLD: Writing, and Speech or Language Disorders (which includes Developmental Speech/Language Disorder, Expressive Language Disorder, Mixed Expressive-Receptive Language Disorder, and Other Speech/Language Disorder).

Results of analyses of the impact of diagnosis on subtest performance indicated that, in many cases, the presence of a diagnosis sometimes indicated poorer performance on tasks than their non-disabled peers. Significant results will be reported here by domain and will be further discussed below in the section titled *Diagnostic Category Impact on Test Performance*. Furthermore, Table 4 summarizes the grand mean scores of each domain in the sample with and without the diagnostic category, which provides additional insight in test performance by diagnosis in the bilingual versus monolingual groups. The grand mean was calculated by taking the average of the reported means for performance on each battery, which allowed for a better understanding of overall performance in the sample by diagnosis.

Fine Motor. Results indicated that individuals who presented with anxiety, speech or language, or specific learning disorder in math diagnoses demonstrated poorer performance on subtests than their non-disabled peers. Those with anxiety performed significantly poorer than their non-disabled peers on Perdue Pegboard assessment of fine motor skills ($t(248) = -2.068$, $p = .040$). Individuals with speech or language diagnoses demonstrated significantly weaker performance than their non-disabled peers on the Grooved Pegboard task ($t(248) = 2.454$, $p = .015$) and fine motor tasks on the NEPSY ($t(109) = 3.323$, $p = .001$). In addition, these individuals performed poorly on tasks of visual-motor integration on the Beery ($t(163) = 3.342$, $p = .001$) and the NEPSY ($t(74) =$

3.040, $p = .003$). Participants who presented with specific learning disorders in math performed significantly poorer than their non-disabled peers on the Grooved Pegboard tasks ($t(248) = 1.949$, $p = .052$). Furthermore, these participants also performed poorly on tasks of visual-motor integration on the Beery ($t(163) = 3.027$, $p = .003$) and NEPSY ($t(74) = 3.227$, $p = .002$).There were no significant differences in performance on tasks of fine motor or visual-motor integration skills in individuals with specific learning disorders in reading or writing or in those with ADHD.

Expressive and Receptive Language. Findings indicated that individuals who presented with anxiety, speech or language disorders, and specific learning disorders in math demonstrated significantly poorer performance on tasks of expressive and receptive language than their non-disabled peers. Individuals with an anxiety disorder performed significantly poorer than their nondisabled peers on expressive ($t(23) = -2.112$, $p = .046$) language tasks on the Clinical Evaluation of Language Fundamentals (CELF). In addition, those with a diagnosis of a speech or language disorder also performed significantly poorer than their non-disabled peers on tasks of expressive and receptive language. Of note, however, is that individuals with a speech or language disorder performed poorly across several batteries, including the CELF ($t(23) = 7.478$, $p = .000$), DKEFS ($t(156) = 5.503$, $p = .000$), and NEPSY ($t(31) = 3.912$, $p = .000$), on expressive language tasks, while their receptive language skills were only significantly poorer than non-disabled peers on the CELF ($t(80) = 9.204$, $p = .000$). When evaluating individuals presenting with a specific learning disorder, it was interesting to note that there were no significant differences in performance in expressive or receptive language among those with reading or writing disabilities. In contrast, those with a math learning disorder performed significantly poorer than their non-disabled peers on expressive language tasks on the CELF ($t(12.050) = 2.567$, $p = .025$) and the DKEFS ($t(156) = 2.701$, $p = .008$) and on receptive language tasks on the CELF ($t(80) = 2.059$, $p = .043$).

Memory. When investigating the impact of presenting diagnoses on memory skills, the following narrow abilities were considered: Immediate Visual Memory, Delayed Visual Memory, Immediate Verbal Memory, Delayed Verbal Memory, and Working Memory. Those who presented with anxiety demonstrated significantly poorer performance on tasks of working memory compared to their non-disabled peers ($t(137) = -1.605$, $p = .028$). Individuals with a speech or language disorder performed significantly weaker than their non-disabled peers on tasks of immediate verbal memory on the CVLT ($t(145) = 3.236$, $p = .001$) and on the WRAML ($t(52) = 4.714$, $p = .000$), tasks of delayed verbal memory on the WRAML ($t(87) = 4.258$, $p = .000$), and tasks of working memory on the WISC ($t(137) = 4.349$, $p = .000$) and Digit Span ($t(122) = 2.367$, $p = .020$). Participants with specific learning disorders in reading demonstrated weak performance on tasks of immediate visual memory on the WRAML ($t(50) = -2.505$, $p = .016$), and working memory tasks on the WISC ($t(137) = 3.100$, $p = .002$) and Digit Span ($t(122) = 2.656$, $p = .009$), when compared to non-disabled peers. Similarly, those with writing learning disorders showed poor performance on tasks of immediate visual memory on the RCFT ($t(250) = 2.470$, $p = .014$), and working memory tasks on the WISC ($t(137) = 2.492$, $p = .014$) and Digit Span ($t(122) = 1.690$, $p = .094$). Finally, individuals with learning disorders in math exhibited significantly poorer performance than non-disabled peers on tasks of delayed verbal memory on the CVLT ($t(250) = 2.309$, $p = .022$) and on working memory tasks on the WISC ($t(137) = 2.959$, $p = .004$).

Long-Term Storage and Retrieval. When evaluating long-term storage and retrieval, the following abilities/domains were considered: Learning Efficiency, Retrieval Fluency, and Speed of Lexical Access. Individuals with Anxiety, ADHD, reading learning disorders, and writing learning disorders demonstrated no significant differences in performance on tasks of long-term memory and retrieval with their non-disabled peers. In contrast, participants with speech or language disorders and math learning disorders demonstrated significant differences in several areas. Individuals with speech or language disorders demonstrated significantly poorer performance on tasks of learning efficiency on the CVLT ($t(145) = 3.853$, $p = .000$) and WRAML ($t(43) = 4.201$, $p = .000$), tasks of retrieval fluency on the NEPSY ($t(31) = 3.912$, $p = .000$) and DKEFS ($t(158) = 4.136$, p .000=), and speed of lexical access on the DKEFS ($t(157) = 5.221$, $p = .000$) and NEPSY ($t(31) = 3.912$, $p = .000$) than their non-disabled peers. Participants with math learning disorders exhibited significantly weaker performance than their non-disabled peers on tasks of retrieval fluency on the DKEFS ($t(158) = 1.995$, $p = .048$), as well as on tasks of speed of lexical access on the DKEFS ($t(157) = 2.161$, $p = .032$).

Executive Functions. Individuals who presented with diagnoses of anxiety, ADHD, and learning disorders in writing demonstrated comparable performance on executive functioning tasks when compared to non-disabled peers. Those presenting with speech or language disorders, reading learning disorders, or math learning disorders, however, demonstrated significant differences in performance on executive functioning tasks as compared to non-disabled peers. Specifically, those with speech or language disorders demonstrated significantly poorer performance than non-disabled peers on tasks of executive functions on the DKEFS ($t(134) = 4.464$, $p = .000$), TOL-2 ($t(178) = 2.507$,

p = .013), and NEPSY (t(106) = 3.413, p = .001). Individuals with learning disorders in reading demonstrated weak performance on executive functioning tasks on the BIBER (t(249) = 2.593, p = .010), NEPSY (t(106) = 2.195, p = .030), and WCST (t(13.510) = 3.380, p = .005). Finally, those with learning disorders in math showed significantly poorer performance than non-disabled peers on executive functioning tasks on the DKEFS (t(134) = 2.324, p = .002) and NEPSY (t(19.731) = 2.187, p = .041).

Cognitive Efficiency. Individuals with anxiety performed significantly poorer than non-disabled peers on cognitive efficiency tasks on the NEPSY (t(26.399) = -2.762, p = .010), Trails (t(123) = -2.535, p = .012), and Digit Span (t(56.65) = -2.135, p = .037). Participants with speech or language disorders exhibited such performance on the NEPSY (t(32) = 3.157, p = .003) and Digit Span (t(122) = 3.035, p = .003). Those with learning disorders in reading showed significantly poorer performance on tasks of cognitive efficiency on the NEPSY (t(32) = 2.623, p = .013) and Digit Span (t(122) = 2.841, p = .005), while those with learning disorders in writing exhibited this weakness on Digit Span (t(18.942) = 2.253, p = .036) and ACT (t(15.436) = 2.649, p = .018). Finally, participants with math learning disabilities showed significantly deficient skills on the Trails task (t(123) = 2.740, p = .007).

Table 4

Grand Means for Domains by Diagnosis

	Grand Mean	
	No Diagnosis	*With Diagnosis*
Fine Motor Domain		
Anxiety	92.35	92.48
ADHD	92.57	92.28
Reading SLD	92.56	92.50
Math SLD	93.23	92.36*
Writing SLD	92.36	92.60
Speech/Language	93.69	86.45*
Language Domain		
Anxiety	102.59	97.06*
ADHD	97.78	101.24
Reading SLD	98.91	95.25
Math SLD	99.10	91.03*
Writing SLD	97.81	98.25
Speech/Language	103.08	84.33*
Long-Term Storage and Retrieval Domain		
Anxiety	98.99	100.91
ADHD	101.42	98.66
Reading SLD	100.02	98.22
Math SLD	99.85	95.76*
Writing SLD	99.34	99.76
Speech/Language	101.47	89.27*
Memory Domain		
Anxiety	95.76	97.06
ADHD	96.18	95.94
Reading SLD	96.30	94.35*
Math SLD	96.56	92.67*
Writing SLD	96.31	94.00*
Speech/Language	97.49	89.64*
Executive Functions Domain		
Anxiety	95.03	97.28
ADHD	96.47	95.05
Reading SLD	96.29	92.06*
Math SLD	95.99	92.41*
Writing SLD	95.67	92.83
Speech/Language	96.34	90.65*

	Grand Mean	
	No Diagnosis	*With Diagnosis*
Cognitive Efficiency Domain		
Anxiety	95.40	94.56*
ADHD	95.00	94.58
Reading SLD	95.67	92.50*
Math SLD	95.26	89.92*
Writing SLD	94.83	91.31*
Speech/Language	95.86	95.55*

An asterisk (*) denotes statistical significance at the p < .05 level on independent samples t-tests performed (discussed in corresponding domain sections).

Chapter VI

Discussion

This study investigated the comparability of the neuropsychological and Cattell-Horn -Carroll (CHC) domains, including the construct equivalence of batteries and tests used regardless of theoretical orientation, the way language influences that comparability and whether monolingual or bilingual status in high SES populations is influencing the constructs being measured. Specifically, this study's intent was to evaluate the following hypotheses: (1) variability in how the neurocognitive domains are represented and constructed in neuropsychological tests and batteries as compared to other similar tests and batteries; (2) the impact of language in the measurement of the other neuropsychological domains, other than measurement of the language domain, according to how much language is used in the measurement of each battery or test; and (3) the impact of the variations in language loadings within each neuropsychological domain, other than in the assessment of the language domain itself, on assessment of bilingual populations compared to monolingual populations.

Results of primary analyses revealed that there existed significant variation in how domains are constructed on certain batteries (Hypothesis 1). In other words, depending upon the test battery used, measurement of the same construct was variable. Specifically, it appears as though there is significant variation in the measurement of fine motor skills on each of the batteries evaluated (Grooved Pegboard, Perdue Pegboard, and NEPSY). In addition, measurement of executive functions on the DKEFS and the Stroop Test revealed significant variation; however, there was no difference noted in the evaluation of other batteries which measure executive functions, such as the Tower of

London, 2nd Edition (TOL-2), Wisconsin Card Sort Test (WCST), NEPSY, and BIBER. Finally, variation was observed in the measurement of cognitive efficiency on the Trails Task, Digit Span Task, and Auditory Consonant Trigrams (ACT). The variation observed suggests that the construct being measured on each battery differs, the implication of which should be considered when selecting test batteries to administer in the assessment of these constructs.

Results of analyses of the second hypothesis revealed no significant differences. In fact, the results obtained suggested that bilingual individuals outperformed monolingual individuals on tests of expressive language. This clearly contradicts the prevailing literature and, as such, led to the question of why these results might occur. Initial hypotheses included that there might be variability in the composite scores created which was creating a narrow range of scores, that the monolingual and bilingual groups were not homogenous, and that there was error in the representation of individuals with speech or language impairments which might be impacting the results. In addition, given that the patient repository was obtained from a private clinic that did not take insurance, it may also be assumed that patients that sought evaluations at this clinic belonged to a higher socio-economic category. This supports that individuals included in this study were largely high performing bilinguals. Given that language was not playing much of a role in the sample, likely due to the impact of high-functioning (high SES) bilinguals, other than perhaps in the manifestations of the different types of disabilities, the scope of this study shifted in focus to examine the impact of presenting diagnosis on neuropsychological test performance in a high SES bilingual population (see section titled: Diagnostic Category Impact on Test Performance).

The results from analyses of the third hypothesis revealed significant differences in several batteries by domain: Delayed Visual Memory as measured on the Rey Complex Figure-Drawing Test (RCFT), Executive Functions as measured by the Tower of London Test, 2nd Edition (TOL-2), and Cognitive Efficiency as measured by Auditory Consonant Trigrams (ACT). These results suggest that the amount of language included (i.e., receptive or expressive language) in the subtests measuring these specific domains on the indicated batteries impacts bilingual individuals' performance on those tests. The impact of language may be present in several ways. This may include language presented in the task directions, language in the item questions, or language required to respond verbally to questions.

Post-hoc analyses attempted to discover the reasons for which the impact of language was not more prominent for bilinguals' performance. As discussed above, these analyses revealed that the sample of bilingual individuals was too homogenous and prevented analyses from identifying significant differences in the impact of language on test performance directly. However, during attempts to group participants by diagnostic category, instead of bilingual or monolingual status, significant differences were found in several areas of performance (see Table 4). The following section will discuss in detail the implications of these results and how they might still point towards, in some ways, the impact of language development overall in test performance among the high SES bilingual population.

Diagnostic Category Impact on Test Performance

In order to facilitate discussion of these results, they will be presented first by domain measured (i.e., Fine Motor skills) and then by the presence or absence of a disorder in each diagnostic category (i.e., ADHD). Reference to Table 4 may be helpful in review of this section. The domain groups to be evaluated are the following: Fine Motor (including Visual-Motor Integration), Expressive/ Receptive Language, Memory (including Visual/Verbal Immediate and Delayed Memory, as well as Working Memory), Long-Term Storage and Retrieval (including Learning Efficiency, Retrieval Fluency, and Speed of Lexical Access), Executive Functions, and Cognitive Efficiency (which shall be considered speed of processing and separate from executive functions). Furthermore, the diagnostic categories to be considered include: Anxiety, ADHD (which includes all subtypes: Inattentive, Hyperactive/ Impulsive, and Combined), Specific Learning Disorder (SLD): Reading, SLD: Math, SLD: Writing, and Speech or Language Disorders (which includes Developmental Speech/Language Disorder, Expressive Language Disorder, Mixed Expressive-Receptive Language Disorder, and Other Speech/Language Disorder).

Fine Motor. Based upon the data available, it appeared that, upon assessment of fine motor functioning, in individuals with an anxiety, speech or language, or specific learning disorder (reading, writing and/or math) diagnosis, there was a significant impact of the presence of a disorder and performance on subtests. Specifically, the sample of individuals who presented with a speech or language disorder and a specific learning disorder in math performed significantly poorer on fine motor tasks overall. Individuals with a speech or language disorder and a specific learning disorder in math diagnosis performed significantly poorer than their non-disabled peers on fine motor tasks as assessed on the Beery, the NEPSY, and Grooved Pegboard. When considering why individuals with these diagnoses would struggle on fine motor tasks, it calls into question an analysis of other task demands, such as the impact of language in the directions or the visual-spatial/orthographic demands (as required in math). Specifically, the presented instructions for all tasks are lengthy and, although they can be repeated, they cannot be reworded to improve an individual's understanding. Further, all tasks require some level of visual-spatial skill (i.e., visual-motor integration tasks on the Beery and NEPSY and the need to rotate the pegs to properly fit the grooves in the slots). Finally, language and fine motor skills are located in the same brain areas and are often seen to be comorbid deficits.

Furthermore, individuals with a diagnosis of anxiety or a specific learning disorder in writing seemed to perform significantly poorer than individuals without a diagnosis when assessed using Purdue Pegboard, but not when assessed using other fine motor batteries. It is important to note that the Purdue Pegboard assessment differs from the Grooved Pegboard in that an individual is asked to place pegs in a straight line down a board, whereas on Grooved Pegboard individuals are asked to securely fit the pegs with grooves in rows and are prompted to continue from left to right. It is interesting, however, that on a seemingly less intense task, individuals with anxiety and writing disorders perform poorly on Purdue Pegboard but not the above-mentioned fine motor assessments. Regardless of understanding the cause of such results, however, awareness of the potential impact on performance for students with these diagnoses is important for practitioners to consider when planning their assessments and/or interpreting their results, as it appears that presenting diagnoses impact performance on these fine motor tasks in significant ways.

Expressive and Receptive Language. When evaluating the impact of presenting diagnoses on language skills, the expressive and receptive language domains were considered. Based upon the data available, it appeared that, upon assessment of expressive and receptive language, in individuals with an anxiety, speech or language, or math learning disorder, there was a significant impact of the presence of a disorder and performance on subtests. Specifically, individuals identified as having an anxiety, speech or language, or math learning disorder diagnosis performed significantly poorer than their non-disabled peers on tests of both expressive and receptive language. Interestingly, there was no such impact of disorder on performance for those with diagnoses of reading or writing learning disorders.

Individuals with an anxiety disorder performed significantly poorer than their nondisabled peers on both expressive and receptive language tasks on the Clinical Evaluation of Language Fundamentals (CELF). Furthermore, those with anxiety disorders struggled more on tasks of expressive language than on those measuring receptive language. This may be the result of task demands to explain, provide definitions, and verbally report responses on expressive language assessments, as opposed to pointing or gesturing on receptive tasks. Interestingly, individuals with anxiety did not perform significantly poorer than their non-disabled peers on language tasks on batteries other than the CELF. This supports that practitioners should use comprehensive assessment strategies and be mindful to use various batteries to assess language and not rely on one battery, as, in this case, individuals with anxiety performed poorer on the CELF than on other batteries (i.e., NEPSY, DKEFS).

In addition, those with a diagnosis of a speech or language disorder also performed significantly poorer than their non-disabled peers on tasks of expressive and receptive language. Of note, however, is that individuals with a speech or language disorder performed poorly across several batteries, including the CELF, DKEFS, and NEPSY, on expressive language tasks, while their receptive language skills were only significantly poorer than non-disabled peers on the CELF. This supports the use of various batteries to evaluate individuals to provide differential information to support or refute the presence of a disorder. This may also suggest that certain batteries are more or less sensitive to impairments.

When evaluating individuals presenting with a specific learning disorder, it was interesting to note that there were no significant differences in performance in expressive or receptive language among those with reading or writing disabilities. In contrast, those with a math learning disorder performed significantly poorer than their non-disabled peers on expressive and receptive tasks. Those with a math disorder performed poorly on expressive language tasks on the CELF and the DKEFS and on receptive language tasks on the CELF. Why would the presence of a math disorder impact performance on expressive and receptive language tasks? Why would individuals with language-based disorders, such as reading and writing, perform no different than their non-disabled peers on expressive and receptive language tasks? These questions are clearly areas for further study and research, which might include questions such as: “What similarities exist between task demands on language subtests and math difficulties?” and “What differences exist on language tasks and reading/writing tasks that differentiate a language disorder versus a learning disorder?”

Memory. When investigating the impact of presenting diagnoses on memory skills, the following narrow abilities were considered: Immediate Visual Memory, Delayed Visual Memory, Immediate Verbal Memory, Delayed Verbal Memory, and Working Memory. Working Memory tasks require an individual to hold simple information, such as numbers or images, in immediate memory and either repeat it back or manipulate it in some way. Interestingly, individuals who presented with any of the diagnosis being evaluated demonstrated statistically significant difficulty with tasks of Working Memory. It seems logical that individuals who struggle with anxiety would perform poorly on these tasks, as they may feel worry or nervousness over remembering the presented items and correctly responding to questions.

Speech, reading, and writing disabilities are all language-based disorders, which are, thus, understandably impacted by the language used in directions and the complexity of the task. Working memory is required for individuals to understand language receptively, to plan expressive language, to read sentences, and to generate ideas into written expression. In addition, individuals with math disorders struggled on tasks of working memory, which points to the need for working memory skills in remembering steps in solving a problem, numbers to borrow, etc. Thus, it seems that working memory is an area largely impacted by presenting disorders and suggests that it may be particularly sensitive to the symptoms associated with these disorders in particular.

For individuals with a speech or language disorder, other areas of memory were also affected and statistically significant, namely immediate and delayed verbal memory. These tasks comprised of list-learning, primarily, which requires an individual to listen to a list of words and repeat them back over a series of trials. This task relies largely on language to present not only directions, but also the list of words to be remembered. In addition to the initial learning task, a delayed task requires the individual to recall and/or recognize the words they learned after a period of time. The demands of this task, linguistically, are nearly identical to the initial task, and it is unsurprising that individuals with speech development difficulties present with weaknesses on this task. This begins to approach the initial hypotheses of this research investigation, in that individuals with weaker English language development would perform more poorly on tasks requiring increased linguistic demands. In addition, this suggests that it would be important to administer these kinds of tasks to understand individual needs as related to speech/language memory; however, it is also imperative to interpret results mindfully, in the context that higher linguistic demands will significantly decrease an individual's performance on tasks intended to measure memory.

Individuals who presented with a specific learning disability in reading also presented with statistically significant difficulties in the areas of delayed verbal memory, as well as immediate visual memory. This pattern is interesting as it seems to incorporate not just the verbal aspects of reading, but also the orthographic (or visual). As discussed above, the presence of a language-based disorder would undoubtedly impact the ability to perform on verbal tasks; however, it is interesting that those with reading disabilities struggled with delayed verbal memory but not immediate verbal memory. This might suggest that this sample of individuals al struggled with reading comprehension and recalling information after it has been read. Furthermore, subjects with reading disabilities also showed significantly weaker immediate visual memory abilities than their peers without diagnoses. This points to the need to understand not only the individual’s ability to understand and recall what they have read, but also their ability to visually represent words in their mind (orthographics), which will impact fluency and, therefore, comprehension. The combination of these results indicates that it would be important to administer assessments in these areas in order to accurately depict the scope of the needs of the individual with the reading disability and plan appropriately for interventions or accommodations to address the memory weaknesses.

Those presenting with specific learning disorders in writing demonstrated statistically significant differences from their peers without a diagnosis in the areas of immediate and delayed visual memory. Writing not only involves the process of planning and organizing thoughts, but also involves fine motor skills, visual-motor integration, and memory skills. For example, individuals must visually represent words, sentences, and paragraphs on paper (orthographics) when writing. Thus, although visual memory may not be an area of assessment that is required for determination of a diagnosis of a writing disorder, these additional assessments may be helpful to administer to students who are presenting with visual-motor, visual-spatial, or orthographic difficulties as part of their writing disorder presentation, as the results yielded can be used to individualize intervention or accommodation plans.

Finally, individuals who presented with a specific learning disorder in math, demonstrated statistically significant weaknesses in delayed verbal memory, as compared to their peers with no disorder. This presentation may be due to comorbidities with other language-based learning disorders (i.e., reading or writing specific learning disorders). It is also possible that individuals with math learning disorders struggle to understand the “language of math” and, therefore, need various repetitions and practice in order to grasp concepts. Within the limited scope of this study and its sample, this may be an area in need of further research.

Long-Term Storage and Retrieval. Long-term Storage and Retrieval, although often considered to be equivalent with "long-term memory," is made up of several other abilities which reflect an individual's ability to properly store information and retrieve it quickly for use in daily like. These abilities include Learning Efficiency, Retrieval Fluency, and Speed of Lexical Access. Learning Efficiency includes the ability to remember previously unrelated information after being paired, remembering semantically related information, and freely recalling information from memory. Retrieval Fluency includes the ability to rapidly access information that is already known or learned and the ability to rapidly produce original thoughts or ideas stemming from that knowledge. Finally, Speed of Lexical Access is the ability to rapidly call objects by their name, identify letters and numbers, and rapidly identify words that are semantically or categorically related (i.e., animals or words that start with the letter "A").

In this study's sample, individuals who presented with a speech or language disorder or a specific learning disorder in math demonstrated statistically significant differences from their peers with no diagnosis on these tasks. Specifically, those with a speech or language disorder performed significantly poorer on all tasks of long-term storage and retrieval (learning efficiency, retrieval fluency, and speed of lexical access); whereas, individuals with a math learning disorder showed poor performance on retrieval fluency and speed of lexical access tasks alone. All these tasks require a significant level of linguistic demand, not only receptively (in order to understand the directions, etc.), but also expressively (i.e., to verbally state recalled information or rapidly name objects). It is not surprising to see, therefore, that those with speech or language disorders perform significantly poorer than their peers on these tasks, as they have existing difficulties with speech (i.e., articulation, dysfluencies, etc.) or receptive/expressive language. As a result, it is important for evaluators to consider that those with suspected (or present) speech or language disorders may perform poorly on tasks of long-term storage and retrieval due to linguistic demands and efforts should be made to interpret results with knowledge of the implications of language demands on the tasks administered. Finally, this also approaches what this study initially intended to investigate, in that, in those with limited language development (disorder or English Language Learner), the impact of linguistic demands of the assessment is likely to affect the individuals' performance more adversely than their monolingual or non-diagnosed peers.

Those presenting with a specific learning disorder in math demonstrated statistically significant weaknesses in the areas of retrieval fluency and speed of lexical access. It is interesting to note that these individuals did not demonstrate any significant differences in learning efficiency, indicating that they are able to learn new information equally well as their non-diagnosed peers. However, their presenting weaknesses fall in line with the symptomology of math disorders; namely, the presence of difficulty with math facts, fluency, and calculations. These results, thus, suggest that it may be helpful to administer these narrow ability subtests during an assessment for a math disability for more information regarding whether the individual has a performance deficit that needs intervention or a cognitive weakness that requires accommodations.

Executive Functions. Evaluation of performance on tasks of executive functioning based upon the presence or absence of a disorder yielded surprising results.

Individuals who presented with diagnoses of anxiety, ADHD, and learning disorders in writing demonstrated comparable performance on executive functioning tasks when compared to non-disabled peers. This seems to be contrary to prevalent research in the area of executive functioning, Anxiety, and ADHD, namely that executive functioning deficits are the hallmark symptoms in individuals with anxiety and ADHD, as well as typical in those with writing learning disorders. The lack of significant differences between those with and without a diagnosis suggests that either the assessments being used are not tapping into the deficits of these individuals (i.e., complex planning, organization, and time management) or that the individuals sampled, given their high SES backgrounds, had access to treatments and supports to address executive functioning weaknesses. Ultimately, these are speculations based upon the results obtained and merit further study and consideration with the current body of research.

Those presenting with speech or language disorders, reading learning disorders, or math learning disorders, however, demonstrated significant differences in performance on executive functioning tasks as compared to non-disabled peers. Specifically, those with speech or language disorders and math learning disorders performed poorly on executive functioning tasks on the DKEFS, TOL-2, and NEPSY, while those with a reading learning disorder performed poorly on the BIBER, NEPSY, and WCST as compared to non-disabled peers.

Cognitive Efficiency. It is unsurprising to note that there is significant impact of speed of processing information (cognitive efficiency) on those with diagnoses of anxiety, ADHD, speech or language disorders, and specific learning disorders. The ability to take information in, analyze and synthesize it quickly, and make speedy decisions is important for all aspects of life and difficulties in one or more areas of development (i.e., emotional, learning) will impact an individual's performance. The findings of this study, thus, confirm the above and, further, provide evidence to support that tasks as basic as Digit Span are consistently effective at evaluating this domain. This study's findings revealed that individuals with anxiety, ADHD, speech or language disorders, and learning disorders in reading, writing, and math all performed significantly poorer on tasks of cognitive efficiency than their non-disabled peers. Cognitive efficiency was measured using the NEPSY, Trail Making Test, Digit Span, and Auditory Consonant Trigrams (ACT). Of note, all individuals with a diagnosis performed significantly poorer than non-disabled peers on the NEPSY, Trail Making Test, and Digit Span. This suggests that Auditory Consonant Trigrams (ACT) is not the best measure of cognitive efficiency.

Study Limitations

There are several limitations of this study which will be considered below. Although efforts were made to address study limitations, no research is without weaknesses. Thus, the following should be considered opportunities for further researchers to expand upon this study's limitations and contribute additional knowledge to the field of psychology, neuropsychology, school psychology, and school neuropsychology.

An important limitation to address is the small sample size. Prior to beginning data collection, preliminary research was conducted to determine the most adequate sample size; however, given the lack of research, this writer and her mentor determined an arbitrary number of 200 subjects to be sufficient. Although this researcher collected over 200 subjects' data, during analyses, many of these subjects were excluded and the projected sample size of 100 subjects per linguistic group was not met. Further research should consider that an increased sample size of over 200 subjects would be beneficial in a similar study.

Due to the lack of available data, namely bilingual subjects, many linguistic analyses could not be run. As such, the study evaluated differences in performance in high SES bilingual and monolingual individuals as moderated by presenting diagnoses on neuropsychological domains. This was not the original intent of the study, which was to evaluate the impact of linguistic demand on bilingual and monolingual individuals' performance on neuropsychological domains; however, valuable information was still gleaned from the data gathered. Future researchers should consider replicating this study's original intent with sufficient subjects in the bilingual and monolingual groups.

In addition, this study's findings supported to a degree that bilingual individuals in high SES households is a protective factor in language development, from parent availability to access to resources. A stronger relationship could have been established in this regard if data was available for individuals from low SES backgrounds. However, given that the study's data was obtained from a private clinic where limited insurance was taken, most individuals were from a high SES background. Further research should consider an analysis of the impact of language and language loading on similar assessment batteries and tests in both high and low SES backgrounds.

Chapter VII

Implications for the Practice of School Psychology

Research would certainly be unnecessary if there were no implications for practitioners in the field! This study found several relevant implications, not just for psychologists or neuropsychologists, or even school neuropsychologists, but for school psychologists as a whole. The results of this study shed light on the need for training programs to emphasize culturally and linguistically diverse practices in assessment. Further, it is imperative that practitioners evaluate an individual as the whole person and identify how their presenting difficulties, family history, personal background, etc. may impact the results obtained on a standardized assessment. The following are implications for psychologists to consider in their professional practices:

1. Awareness of the potential impact on performance for students with these diagnoses is important for practitioners to consider when planning their assessments and/or interpreting their results, as it appears that presenting diagnoses impact performance tasks in significant ways.
2. Awareness of the presenting differences in SES, family backgrounds, etc. that impact an individual's skills and experiences. For example, this study's findings included that higher SES bilinguals appear to be higher functioning and high SES may be a protective factor in language development.
3. Practitioners should use comprehensive assessment strategies and be mindful to use various batteries to assess language (and other domains). In addition, evaluations should attempt to differentially diagnose individuals to support or refute the presence of a disorder. It is imperative to conduct comprehensive

assessments with more than one battery, as this study found that not all batteries measure the same construct in the same way.

Areas for Further Study and Future Research

In addition to the above implications for the practice of school psychology, questions were raised that were beyond the scope of this study and warrant further research. These topics and/or questions are listed here as areas for further study.

1. Why would the presence of a math disorder impact performance on expressive and receptive language tasks? Why would individuals with language-based disorders, such as reading and writing, perform no different than their non-disabled peers on expressive and receptive language tasks? These questions are clearly areas for further study and research, which might include questions such as: “What similarities exist between task demands on language subtests and math difficulties?” and “What differences exist on language tasks and reading/writing tasks that differentiate a language disorder versus a learning disorder?”
2. Further investigation into assessment of high SES bilinguals versus low SES bilinguals and the impact of language development on their performance on neuropsychological batteries, with a particular focus on the possibility of high SES as a protective factor for language development.
3. Further research into neuropsychological assessment of bilinguals versus monolinguals and the impact of language.

References

Ardila, A. (1992). Luria's approach to neuropsychological assessment. *International Journal of Neuroscience, 66,* 35 – 43. doi:10.3109/00207459208999787

Bialystok, E. (2011). Coordination of executive functions in monolingual and bilingual children. *Journal of experimental child psychology*, *110*(3), 461-468. doi10.1016/j.jecp.2011.05.005

Bialystok, E. (2010). Global-local and trail-making tasks by monolingual and bilingual children: Beyond inhibition. *Developmental Psychology, 46,* 93-105. doi:10.1037/a0015466

Bialystok, E. (2001a). *Bilingualism in development: Language, literacy, and cognition*. New York, NY: Cambridge University Press.

Bialystok, E. (2001b). Metalinguistic aspects of bilingual processing. *Annual Review of Applied Linguistics*, *21*, 169-181. doi:10.1017/S0267190501000101

Bialystok, E. (1999). Cognitive complexity and attentional control in the bilingual mind. *Child Development, 70,* 636-644. doi:10.1111/1467-8624.00046

Bialystok, E., & Barac, R. (2012). Emerging bilingualism: Dissociating advantages for metalinguistic awareness and executive control. *Cognition*, *122*(1), 67-73. doi:10.1016/j.cognition.2011.08.003

Bialystok, E., & Viswanathan, M. (2009). Components of executive control with advantages for bilingual children in two cultures. *Cognition*, *112*(3), 494-500. doi:10.1016/j.cognition.2009.06.014

Bialystok, E., Barac, R., Blaye, A., & Poulin-Dubois, D. (2010). Word mapping and executive functioning in young monolingual and bilingual children. *Journal of Cognition and Development*, *11*(4), 485-508. doi:10.1080/15248372.2010.516420

Bialystok, E., Craik, F I. M., & Luk, G. (2008). Cognitive control and lexical access in younger and older bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34,* 859-873. doi:10.1037/0278-7393.34.4.859

Bialystok, E., Craik, F 1. M., & Luk, G. (2012). Bilingualism: Consequences for mind and brain. *Trends in Cognitive Sciences, 16,* 240-250. doi:10.1016/j.tics.2012.03.001

Bialystok, E., Craik, F 1. M., & Ruocco, A. C. (2006). Dual-modality monitoring in a classification task: The effects of bilingualism and aging. *Quarterly Journal of Experimental Psychology, 59,* 1968-1983. doi:10.1080/17470210500482955

Bialystok, E., & Martin, M. M. (2004). Attention and inhibition in bilingual children: Evidence from the dimensional change card sort task. *Developmental Science, 7,* 325-339. doi:10.1111/j.1467-7687.2004.00351.x

Carlson, S. M., & Meltzoff, A. N. (2008). Bilingual experience and executive functioning in young children. *Developmental Science, 11,* 282-298. doi:10.1111/j.1467-7687.2008.00675.x

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge, England: Cambridge University Press.

Cattell, R. B. (1971). *Abilities: Their structure, growth, and action.* Boston, MA: Houghton Mifflin.

Collier, V. P. (1995). Acquiring a second language for school. *Directions in Bilingual Education, National Clearinghouse for Bilingual Education, 1,* 1-8. Retrieved from http://files.eric.ed.gov/fulltext/ED394301.pdf

Cormier, D. C., McGrew, K. S., & Ysseldyke, J. E. (2014). The influences of linguistic demand and cultural loading on cognitive test scores. *Journal of Psychoeducational Assessment, 32*. 610-623. doi:10.1177/0734282914536012

Costa, A., Hernández, M., Costa-Faidella, J., & Sebastián-Gallés, N. (2009). On the bilingual advantage in conflict processing: Now you see it, now you don't. *Cognition*, *113*(2), 135-149. doi:10.1016/j.cognition.2009.08.001

Cummins, J. C. (1984). Bilingual and special education: Issues in assessment and pedagogy. Austin, TX: PRO-ED. In Flanagan, D. P., Ortiz, S. O., & Alfonso, V. C. (2013). *Essentials of cross-battery assessment* (3rd ed.). Hoboken, NJ: John Wiley and Sons, Inc.

Feifer, S. G., & DeFina, P. D. (2002). *The neuropsychology of written language disorders: Diagnosis and intervention*. Middletown, MD: School Neuropsych Press.

Feifer, S. G. & DeFina, P. D. (2000). *The neuropsychology of reading disorders: Diagnosis and intervention.* Middletown, MD: School Neuropsych Press.

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

Flanagan, D. P., Alfonso, V. C., Ortiz, S. O., & Dynda, A. M. (2008). Best practices in cognitive assessment. In A. Thomas & J. Grimes (Eds.) *Best practices in school psychology V* (pp. 1 – 27). Washington, DC: National Association of School Psychologists.

Flanagan, D. P., & McGrew, K. S. (1998). Interpreting intelligence tests from contemporary Gf-Gc theory: Joint confirmatory factor analysis of the WJ-R and KAIT in a non-white sample. *Journal of School Psychology*, *36*(2), 151-182. doi:10.1016/S0022-4405(98)00003-X

Flanagan, D. P., McGrew, K. S., & Ortiz, S. O. (2000). *The Wechsler Intelligence Scales and Gf-Gc theory: A contemporary approach to interpretation*. Needham Heights, MA: Allyn & Bacon.

Flanagan, D. P. & Ortiz, S.O. (2007). *Essentials of Cross-Battery Assessment* (2nd ed.). New York, NY: John Wiley and Sons, Inc.

Flanagan, D. P., Ortiz, S. O., & Alfonso, V. C. (2013). *Essentials of Cross Battery Assessment* (3rd ed.). New York, NY: John Wiley and Sons, Inc.

Flanagan, D. P., Ortiz, S. O., Alfonso, V. C., & Mascolo, J. T. (2006). Academic Achievement and the Cattell-Horn-Carroll (CHC) Theory. In Flanagan, D. P. (Ed.). *Achievement test desk reference (ATDR): A guide to learning disability identification* (2nd ed.). (pp. 18 – 47). Hoboken, NJ: John Wiley and Sons, Inc.

Foy, J. G., & Mann, V. A. (2014). Bilingual children show advantages in nonverbal auditory executive function task. *International Journal of Bilingualism*, *18*(6), 717-729. doi:10.1177/1367006912472263

Horn, J. L. & Noll, J. (1997). Human cognitive capabilities: *Gf-Gc* theory. In D. P. Flanagan, J. L. Genshaft, & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 53 – 91). New York, NY: Guilford Press.

Kovács, Á. M., & Mehler, J. (2009). Cognitive gains in 7-month-old bilingual infants. *Proceedings of the National Academy of Sciences*, *106*(16), 6556-6560. doi:10.1073pnas.0811323106

Kranzler, J. H., Flores, C. G., & Coady, M. (2010). Examination of the cross-battery approach for the cognitive assessment of children and youth from diverse linguistic and cultural backgrounds. *School Psychology Review, 39,* 431-446. Retrieved from https://www.researchgate.net/profile/Maria_Coady/publication/228497938_Examination_of_the_crossbattery_approach_for_the_cognitive_assessment_of_children_and_youth_from_diverse_linguistic_and_cultural_backgrounds/links/557c4a3508ae26eada8c9c81.pdf

Kroll, J. F., and Bialystok, E. (2013). Understanding the consequences of bilingualism for language processing and cognition. *Journal of Cognitive Psychol*ogy 25, 1–18. doi:10.1080/20445911.2013.799170

Luria, A. R. (1980). *Higher cortical functions in man* (2nd ed.). New York: Basic Books.

Luria, A. R. (1973). *The working brain.* New York: Basic Books.

Luria, A. R. (1966). *Human brain and psychological processes.* New York: Harper and Row.

Martin-Rhee, M. M., & Bialystok, E. (2008). The development of two types of inhibitory control in monolingual and bilingual children. *Bilingualism: language and cognition*, *11*(01), 81-93. doi:10.1017/S1366728907003227

McGrew, K. S., & Flanagan, D. P. (1998). *The intelligence test desk reference (ITDR): Gf-Gc cross-battery assessment*. Needham, MA: Allyn & Bacon.

Miller, D. C. (2015, September). Integration of the WJ IV, WISC-V, WISC-V Integrated, and WIAT-III into a School Neuropsychological Assessment Model. In *School Psychology Forum* (Vol. 9, No. 3). Retrieved from https://www.nasponline.org/publications/periodicals/spf/volume-9/volume-9-issue-3-(fall-2015.

Miller, D. C. (Ed). (2013). *Best practices in school neuropsychological assessment.* Hoboken, NJ: John Wiley and Sons, Inc.

Morales, J., Gómez-Ariza, C. J., & Bajo, M. T. (2013). Dual mechanisms of cognitive control in bilinguals and monolinguals. *Journal of Cognitive Psychology*, *25*(5), 531-546. doi:10.1080/20445911.2013.807812

Morton, J. B. (2010). Language, bilingualism, and executive functioning in early development. *Psychological reports*, *107*(3), 888-890. doi:10.2466/04.11.28.PR0.107.6.888-890

National Association of School Psychologists. (2015). *The provision of school psychological services to bilingual students* [Position statement]. Bethesda, MD.

National Center for Education Statistics. (2013). *The nation's report card: Trends in academic progress 2012* (NCES Publication No. 2013456). Washington, DC: U.S. Department of Education, Institute of Education Sciences. Retrieved from https://nces.ed.gov/nationsreportcard/pubs/main2012/2013456.aspx

National Center for Education Statistics. (2016). *The condition of education: English language learners in public schools*. Washington, D.C.: U.S. Department of Education, Institute of Education Sciences. Retrieved from https://nces.ed.gov/programs/coe/indicator_cgf.asp

National Center for Education Statistics. (2017). *The condition of education: English language learners in public schools*. Washington, D.C.: U.S. Department of Education, Institute of Education Sciences. Retrieved from https://nces.ed.gov/programs/raceindicators/indicator_RBC.asp

Ortiz, J. A., Devine, R. I., & Ortiz, S. O. (2016). Assessment of culturally and linguistically diverse individuals with the WJ-IV. In F. Schrank, S. Decker, & J. Garruto (Eds.) *Essentials of WJ IV Cognitive Abilities Assessment*. Hoboken, NJ: John Wiley and Sons, Inc.

Ortiz, S. O., Melo, K., Terzulli, M. A. (2017). Assessment of culturally and linguistically diverse individuals with the WISC-V. In D. Flanagan & V. Alfonso (Eds.) *Essentials of WISC-V Assessment*. Hoboken, NJ: John Wiley and Sons, Inc.

Ortiz, S. O. (2016, July). *Neuropsychological evaluations of English language learners: Best practices in nondiscriminatory assessment of linguistically and culturally diverse individuals.* Lecture presented at School Neuropsychology Summer Institute, Texas.

Ortiz, S. O. (2011) Separating cultural and linguistic differences (CLD) from specific learning disability (SLD) in the evaluation of diverse students: Difference or disorder? In D. P. Flanagan & V. C. Alfonso (eds.). *Essentials of Specific*

Learning Disability Identification (pp. 299-325). Hoboken, NJ: John Wiley and Sons, Inc.

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

Ortiz, S. O., Flanagan, D. P., & Alfonso, V. C. (2015). Culture-language interpretive matrix. On *Cross-battery assessment software system (X-BASS v1.0) (CD-ROM).* Hoboken, NJ: John Wiley and Sons, Inc.

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. 1 – 18). Washington, DC: National Association of School Psychologists.

Paap, K. R., & Sawi, O. (2014). Bilingual advantages in executive functioning: problems in convergent validity, discriminant validity, and the identification of the theoretical constructs. *Frontiers in psychology*, *5*, 962. doi:10.3389/fpsyg.2014.00962

Rhodes, R., Ochoa, S. H. & Ortiz, S. O. (2005). Comprehensive Assessment of Culturally and Linguistically Diverse Students: A practical approach. New York: Guilford.

Sattler, J. M. (2001). *Assessment of children: Behavioral and clinical applications* (4th ed.). San Diego, CA: Jerome M. Sattler.

Sotelo-Dynega, M., Ortiz, S. O., Flanagan, D. P. & Chaplin, W. (2013). English Language Proficiency and Test Performance: Evaluation of bilinguals with the Woodcock-Johnson III Tests of Cognitive Ability. Psychology in the Schools, Vol 50(8), 781-797.

Templeton, M. M. (2012). *An examination of the effects of culture and language on the executive functioning of Spanish-speaking English learners according to the Delis-Kaplan executive function system.* Doctoral Dissertation. Alliant International University, San Diego, CA.

Thomas, W. P., & Collier, V. P. (2002). *A national study of school effectiveness for language minority students' long-term academic achievement* (Report No. ED475048). Santa Cruz, CA: Center for Research on Education, Diversity, and Excellence.

Valdes, G. and Figueroa, R. A. (1994). Bilingualism and testing: A special case of bias. Norwood, NJ: Ablex Publishing Corporation.

VITA

Name	*Meghan Ashley Terzulli*
Baccalaureate Degree	*Bachelor of Arts*
	St. John's University
	Jamaica, NY
	Major: Psychology
Date Graduated	*May, 2013*
Other Degrees	*Master of Science*
	St. John's University
	Jamaica, NY
	Major: School Psychology,
	Bilingual Track
Date Graduated	*May, 2016*