12-2016

False positives on neuropsychological measures of effort among bilingual neurologically intact Mexican Americans

Arnoldo Amador Jr.
The University of Texas Rio Grande Valley

FALSE POSITIVES ON NEUROPSYCHOLOGICAL MEASURES OF EFFORT AMONG BILINGUAL NEUROLOGICALLY INTACT MEXICAN AMERICANS

A Thesis

by

ARNOLDO AMADOR JR.

Submitted to the Graduate College of
The University of Texas Rio Grande Valley
In partial fulfillment of the requirements for the degree of

MASTER OF ARTS

December 2016

Major Subject: Experimental Psychology

FALSE POSITIVES ON NEUROPSYCHOLOGICAL MEASURES OF EFFORT AMONG

BILINGUAL NEUROLOGICALLY INTACT MEXICAN AMERICANS

A Thesis
by
ARNOLDO AMADOR JR.

COMMITTEE MEMBERS

Dr. Philip Gasquoine
Chair of Committee

Dr. Amy A. Weimer
Committee Member

Dr. Ralph Carlson
Committee Member

December 2016

ABSTRACT

Amador, Arnoldo, <u>False Positives on Neuropsychological Measures of Effort Among Bilingual, Neurologically Intact, Mexican Americans</u>. Master of Arts (MA), December, 2016, 35 pp., 5 tables, references, 52 titles.

The present study examines the performance of bilingual Mexican Americans on neuropsychological measures of effort in language and visual-perceptual formats. Optimal/suboptimal effort cutoff scores derived from monolingual English Speakers were used to test participants divided by language of administration and bilingual groupings based upon Spanish-English difference scores on the Woodcock Munoz Language Survey-Revised Picture Vocabulary subtest. Participants produced more false positives (misidentification of suboptimal effort) on the Reliable Digit Span, a language formatted measure, and similar rates of false positives on visual-perceptual effort measures: Test of Memory Malingering and the Dot Counting Test as compared to established rates using monolingual English speakers. Results show that the Test of Memory Malingering and Dot Counting Test are appropriate for this linguistically diverse population while the Reliable Digit Span is not. These results show why linguistically diverse groups can be misdiagnosed as giving poor effort when the tests being used are not appropriate.

DEDICATION

This goes out to my mother, for without her I would not be the man I am today. A whole hearted and earnest dedication to both my parents, Arnoldo and Maria Amador, for always supporting my dreams and showing me that excellence is a choice, that hard-work is a calling and that without perseverance life would be taken for granted. To all the fires that we have walked through, thank you for strengthening my resolve. Here is to life, the most nurturing and difficult of teachers, for bringing the most wonderful friends and family along for this journey. This is the culmination of experience, the triumph of hardships and the lamentation of the past that has molded us into the present versions of ourselves. Thank you to everyone who has had a hand in this and thank you to my former self, for not giving up no matter the occasion.

This right here would not be possible without the opportunities and guidance I received by my mentors at the University of Texas Rio Grande Valley. Thank you Dr. Rampersad and HHMI, for believing in me and raising the bar for myself and those around. Thank you to Dr. Gasquoine, challenging me and demanding excellence, for always pushing me to do better and better. Thank you to Dr. Weimer, for always going above and beyond for us and for being the glue that kept the lab united. Thank you to Dr. Carlson, for the opportunity to work and learn from you the past two years, they've been full of wonderful stories and lessons that I will always cherish.

ACKNOWLEDGEMENTS

I would like to acknowledge the contributions that made the completion of my thesis possible. The Science Education Grant funded by the Howard Hughes Medical Institute (HHMI) for giving me my first real job and being very generous throughout the entire process. The Center of Excellence in STEM Education for housing me while I worked and my coworkers for being very welcoming and loving. The Center for Bilingual Studies for being another place for me to work and meet wonderful people. I would like to acknowledge the wonderful research assistants and mentors that have made love every single moment of our lab experience. Dr. James Aldridge, for calling me that fateful day in January of 2014 and starting me down this path. Sarah Stegall, for being the first person to show me the tricks of the trade and getting me involved in more research. Yajaira Rivera for always making sure that the Center was always available for us all and for all our conversations over the years. Many research assistants from this lab have made lasting impacts on my life and consider them great friends: Adrianna Sarmiento, Lucila Lopez, Liliana Flores, Cindy Gonzalez, Francisco Sierra, Marlene Orta, David Cruz, Mario Arredondo, Melissa Leon, Daniela Kuri, Yuridia Munoz, Melissa Garza, Alejandra Ballesteros, Deanna Pollard, Luis Perez, Nina Lopez, Linda Torres, and Isela Stephens. You all are such wonderful people that truly have a special place in my heart.

TABLE OF CONTENTS

Page

ABSTRACT iii

DEDICATION iv

ACKNOWLEDGEMENTS v

TABLE OF CONTENTS vi

LIST OF TABLES vii

CHAPTER I. INTRODUCTION 1

CHAPTER II. REVIEW OF LITERATURE: BILINGUALISM 5

Statement of Purpose 10

Hypotheses 10

Hypothesis 1 10

Hypothesis 2 11

Hypothesis 3 11

CHAPTER III. METHODOLOGY 12

Method 12

Participants 12

Measures 12

CHAPTER IV. RESULTS 17

Hypothesis Testing 18

Hypothesis 1 18

Hypothesis 2 20

Hypothesis 3 20

CHAPTER V. DISCUSSION 24

REFERENCES 30

BIOGRAPHICAL SKETCH 35

LIST OF TABLES

Page

Table 1: Means (SDs) for the Test of Memory Malingering (TOMM) Trial 2, Reliable Digit Span, Dot Counting, Woodcock Munoz Language Survey-Revised (WMLS-R) English Score, WMLS-R Spanish score, and degree of bilingualism………18

Table 2: Total Number of False Positives on the Reliable Digit Span between English vs. Spanish Language of Administration………..…………………………………………..19

Table 3: False Positives on the Reliable Digit Span between Balanced and Language-Dominant Bilinguals Tested in English……….…………………………………………19

Table 4: False Positives on the Reliable Digit Span between Balanced and Language-Dominant Bilinguals Tested in Spanish……………………………………………….20

Table 5: Intercorrelations among Age, Education, Income, Test Of Memory Malingering (TOMM) Trial 2, Reliable Digit Span, Dot Counting, Woodcock Munoz Language Survey Revised (WMLS-R) English Score, WMLS-R Spanish Score, and degree of Bilingualism...21

CHAPTER I

INTRODUCTION

Neuropsychological testing is used to determine the presence of neurocognitive impairment. There are assessments that have been developed by neuropsychologists to detect and measure neurocognitive impairment across the five domains of neurocognition (memory, visual-perceptual, language, somatosensory and motor skills, and executive functioning). For example, an assessment such as the Continuous Visual Memory Tests (Trahan & Larrabee, 1988) uses complex and ambiguous designs in a recognition format to measure visual learning and memory. The Continuous Performance Task (Conners et al., 2000) is a computerized assessment that measures participants' inattentiveness, impulsivity, sustained attention and vigilance, all aspects of executive function. The California Verbal Learning Test-II (Delis, Kramer, Kaplan, & Ober, 2000) is one of the most widely used neuropsychological assessments that measures explicit verbal memory by asking participants to recall a list of nouns presented over five learning trials. The Wisconsin Card Sorting Test (Heaton, Chelune, Talley, Kay, & Curtiss, 1993) measures executive functioning in participants by requiring them to sort cards with symbols that vary in color, number, and shape according to different testing principles.

In the event that results on these neuropsychological tests are in the impaired range, explanations aside from neurocognitive impairment must be ruled out. One explanation for poor performance is suboptimal cooperation and effort (Larrabee, 1992). In order for neuropsychological assessments to yield valid results, patients must put forth optimal effort, as lower effort resulting in poor neuropsychological test scores can be misinterpreted by professionals as neurocognitive impairment. Such misinterpretation of results that indicate neurocognitive impairment when a patient is actually neurocognitively intact is called a false positive (Gasquoine & Gonzalez, 2012). Thus, Symptom Validity Testing (SVT) was created in order to assess patient effort (Gasquoine, 2013).

On the surface, SVT measures appear to be high in their difficulty level, but they actually have a low level of difficulty. It is a popular misconception that head trauma/injury causes neurocognitive debilitation when in fact this is the case only for severe injuries. It is such misconceptions that encourage malingerers (people who intentionally fake their symptoms) to exaggerate their neurocognitive impairment (Bianchini, Mathias, & Greve, 2001; Boone et al., 2000; Inman & Berry, 2002). SVTs assess effort, as they reveal patterns or symptoms that suggest participants are deliberately performing poorly (McMillan et al., 2009). There are two types of SVT measures that test for optimal/suboptimal effort: (a) stand-alone measures that were designed specifically to identify poor effort; and (b) embedded measures that are derived from the standard administration of neuropsychological tests (Gasquoine, 2013). Both types of measures can quantitatively account for the amount of effort that a patient is putting forth (Armistead-Jehle & Hansen, 2011).

The Test of Memory Malingering (TOMM: Tombaugh, 1996) is one of the most popular stand-alone SVTs along with some of the older tests such as the Dot Counting Test (Rey, 1941) and Rey 15-Item Memorization Test (Rey, 1964). The TOMM was designed for adults to distinguish between malingerers (people who are intentionally feigning symptoms) and patients with actual memory impairment. In order to quantify patient effort on neuropsychological tests, criterion cutoff scores are used that create an artificial dichotomy of optimal/suboptimal effort

(Gasquoine, 2013). Through validation studies, the suggested criterion cutoff for suboptimal effort with the TOMM is either: (a) ≤ 90% correct on the second trial or the retention trial; or (b) a score lower than chance performance (i.e., < 50%) on any of the three trials (Tombaugh, 1996). This optimal/suboptimal cutoff score was derived through the study of the distribution of test scores from various samples, such as intentional and instructed malingers. In order to properly recognize suboptimal effort, patterns in patient performance are considered (Larrabee, 2003).

The Dot Counting Test was created in the 1940s and it still remains a popular stand-alone effort measure. It uses twelve cards with visually represented dots. The first six cards use an ungrouped arrangement of 7 to 27 dots while the last six use a grouped arrangement of 8 to 28 dots that the patient must count as quickly as possible. A combination score (E-score) is derived from the average time for grouped plus ungrouped dots in seconds plus the total number of errors. The recommended E-score to identify suboptimal effort is ≥ 14 (Boone, Lu, & Herzberg, 2002).

In addition to stand-alone SVTs, there are also embedded SVTs, that are derived from traditional neuropsychological tests used to measure neurocognitive functioning, such as the Wechsler Adult Intelligence Scale- Fourth Edition (WAIS-IV: Wechsler, 2008). Like stand-alone measures, embedded measures assess the validity of patient effort during test performance (Meyers, Volbrecht, Axelrod, & Reinsch-Boothby, 2011). The Reliable Digit Span, (Greiffenstein, Baker, & Gola, 1994) for example, is one of the most widely accepted embedded SVT measures. Recent literature states that the Reliable Digit Span is an effective embedded measure that can detect suboptimal effort in participants (Axelrod, Fichtenberg, Millis, & Wertheimer 2006; Loughan, Perna, & Hertza, 2012). The covert nature of embedded assessments offers an advantage over stand-alone measures. As a result, they can enhance the usage of stand-alone SVTs and can be administered concurrently to best identify the effort a patient is putting forth (Novitski, Steele, Karantzoulis, & Randolph, 2012; Paulson, Horner, & Bachman, 2015).

Studies involving neuropsychological testing are published with the assumption that participants give forth their best effort, thus accurately reflecting their neurocognitive function and capabilities. In the case that patient effort is not optimal, results become skewed and therefore invalid because they did not measure what they originally intended. Underestimating patients' abilities as a result of diagnostic errors calls for the usage of various measures of effort in order to validate a patient's performance by reinforcing sensitivity to suboptimal effort (Heilbronner et al., 2009; Slick, Sherman, & Iverson, 1999). A false positive on an SVT misidentifies a patient as deliberately exaggerating their symptoms, therefore not putting forth their best effort, when in fact they actually are trying. False positives on neurocognitive assessments used to measure neurocognition misidentify a patient as cognitively impaired when they are actually cognitively intact.

Cases of suboptimal effort mostly involve litigation, in which compensation for personal or work-related injury is being sought, or criminal prosecution. Therefore, it is important to acknowledge any potential reinforces for a patient undergoing neuropsychological assessment that can lead to poor performance from the intentional feigning of symptoms (Vickery, Berry, Inman, Harris, & Ore, 2001). The creation of SVTs has helped identify suboptimal effort because of research that establishes their sensitivity (proper identification of suboptimal effort) and specificity (proper identification of optimal effort). Sensitivity (100% – % false negatives) and specificity (100% - % false positives) rates are derived from different types of research studies. For example, when testing for sensitivity, a sample of patients with some form of brain injury are used, while testing for specificity, neurocognitively intact patients are used.

CHAPTER II

REVIEW OF LITERATURE: BILINGUALISM

Neuropsychology in Spanish-speaking Latin American countries has developed at a slower rate than in the United States and Europe. According to Pontón and Ardila (1999), this slower rate can be attributed to the following observations: (a) Latin America is made up of many nations that vary in their scientific and economic development; thus availability of resources is different across the board; (b) intercommunication among countries varies because the means of communication are not yet as up to date as in other growing countries; (c) the economy is very different in Latin American countries as opposed to the United States, wherein private practice and education for neuropsychologists is not as compensatory. As a result, access to high quality journals and other types of publications is limited; and lastly (d) literature available to neuropsychologists in their native language is also difficult to access due to lack of funds.

The method in which neuropsychological assessments are created and administered can affect the delivery of service. More specifically, it is important to include diverse populations when validating assessments so that results can be generalized broadly, interpreted appropriately, and professionals can treat patients accordingly. It ensures that patients receive the appropriate care and diagnoses (Arnold, Montgomery, Castaneda, & Longoria, 1994). The need for culturally aware/competent mental health professionals has grown in recent years (Mungas, Reed, Haan, & González, 2005; Pérez-Arce, 1999). Fortunately, methodological research on the evaluation of neuropsychological assessments for ethnic minorities also has increased. This is important because results on neuropsychological assessments and other derivatives, such as stand-alone and embedded measures of effort, depend on participants being able to fully understand test instructions. However, despite the abundance of neuropsychological test data on monolingual English-speakers that demonstrates better performance compared to ethnic and linguistic minorities (Vilar-Lopéz & Puente, 2010), this area of research frequently does not include Hispanic Americans.

Limited research has examined the validity of stand-alone SVTs such as the TOMM and Dot Counting Test in Spain with Spanish-speakers, wherein their performance was compared to that of monolingual English-speakers (Vilar-López, Gómez-Río, Caracuel-Romero et al., 2008; Vilar-López, Gómez-Río, Santiago-Ramajo et al., 2008). They found the usage of both measures to be effective in distinguishing between malingerers and non-malingers in the Spanish language. However, the TOMM had the better rate of classification between the three groups involved in the study: (a) non-compensation seeking; (b) compensation seeking- not suspected of malingering; and (c) compensation seeking- suspected of malingering.

Cross-language validation of measures presents challenges, including: (a) the translation of measures; (b) normative based interpretation of scores; and (c) the evaluator's competency in the language of administration (Elbulok-Charcape, Rabin, Spandaccini, & Barr, 2014). The translation of tests can appear to be the same but the connotation does not necessarily follow suit because the translated version might not be culturally appropriate (Pontón & Ardila, 1999). Furthermore, cultural awareness and sensitivity are factors that need to be taken into consideration when developing new measures. Cultural biases might complicate the validation of measures for a minority population (Arnold et al., 1994). Testing culturally diverse populations must be approached with caution because data that validates their use may not exist for all populations. For example, assessments that are created in the U.S reflect the cultural bias of who it was intended for, and as a result might be inappropriate for other populations (Elbulok-Charcape et al., 2014).

Sociopolitical contexts of a majority population might occlude those of ethnic minorities, whom are affected by unique contextual factors. The need to understand different populations has been stressed in a socio-political context, but it also needs to be backed by scientific inquiry (Puente & Perez-Garcia, 2000). Such has been the case for Hispanic American ethnic groups, where neurologically intact participants score poorly on neuropsychological tests that were normed for monolingual English-speakers and as a result are misdiagnosed with neurocognitive impairment. Misdiagnoses of understudied ethnic groups further exemplifies the need for more research to be dedicated to these populations to examine the reliability and validity of measures across ethnic groups.

Cutoff scores derived from the distribution of test scores across many sample populations have been created by professionals based on research with neuropsychological assessments. These cutoff scores and their many implications (neurocognitively intact/impaired) need to be validated for their usage on diverse linguistic groups (Puente & Perez-Garcia, 2000). Furthermore, cutoffs should be used with caution because incorrect interpretations can contribute to subsequent stereotypes that suggest possible genetic and or environmental limitations in ethnic minorities (Elbulok-Charcape et al., 2014).

Not many neuropsychological assessments have been normed for use with Hispanic Americans. Consequently, it is common for Hispanic Americans to be misdiagnosed because professionals are using guidelines created for use with the majority population (Vilar-Lopéz & Puente, 2010). One of the few assessment instruments to be validated in the U.S. for use with Spanish-speakers is the Woodcock-Munoz Language Survey-Revised (WMLS-R: Woodcock, Munoz-Sandoval, Reuf, & Alvarado, 2005).

Neuropsychological test norms developed for monolingual English-speakers may not be valid for Hispanic Americans, who might be bilingual. Additionally, the WMLS-R has only been validated for use with monolingual Spanish speakers, leaving questions about its appropriateness for use with bilinguals (Gasquoine, Croyle, Cavazos-Gonzalez, & Sandoval, 2007). According to Vilar-Lopez & Puente (2010), there is a lack of research on the association between bilingualism and neurocognitive capacity, though some research has been done.

Bialystok, Luk, and Kwan (2005) found that bilinguals perform better as compared to monolinguals in visual-perceptual tasks that require response inhibition (suppressing an inappropriate action with regards to a specific task or goal). This reinforced the notion that bilinguals have greater executive functioning when compared to monolinguals (Bialystok, 1999). This claim was corroborated by findings that bilingualism positively enhances attention in the presence of distractors or conflicting information (measures of executive functioning). For example, when compared to monolinguals, bilinguals demonstrated faster recognition on the flanker task, a task that requires the suppression of conflicting stimuli (Costa, Hernandez, Costa-Faidella, & Sebastian-Galles, 2009). As bilinguals use two languages, they must inhibit one vocabulary constantly, when using the other (Green, 1998). This is thought to result in better inhibitory control, as evidenced on tests that require the inhibition of stimuli that may cause interference (Bialystok, 1999). In addition to differentiating between discrepancies, as a result of their bilingualism, they can also go back and forth between languages. Therefore, they also prove more adept at conflict resolution when exposed to different types of congruency trials (ungrouped/grouped; mixed/unmixed).

Compared to monolinguals, bilinguals have unique language processing skills. Bilinguals' vocabulary, when measured in just one of their two languages, is lower than monolinguals (Bialystok, 2009). Overall though, bilinguals have a more expansive vocabulary (if measuring vocabulary additively across the two languages) when compared to monolinguals. As a result, the likelihood of a prolonged delay in access and retrieval of words from their lexical banks is increased (Mindt et. al., 2008). Early in childhood, vocabulary is oftentimes used to measure how well children may or may not understand language.

Because bilingual children have two languages to learn they may not show proficiency in one language until a little further in development as compared to monolinguals who only have one language (Bialystok, 2009; Mindt et. al., 2008). Furthermore, adult bilinguals access and retrieve their vocabularies, perhaps because of its expansiveness, more gradually as compared to monolinguals who do not have lexical banks in two languages. Bilinguals receive lower scores, when compared to monolinguals on language tasks in which participants are asked to name pictures only in one language while inhibiting the other (Mindt et. al., 2008). Another example can be derived from mean scores of Hispanic Americans on the Wechsler Intelligence Scale for Children, 3rd edition (WISC-III: Wechsler, 1991), where Verbal IQ (but not Performance IQ) scores were approximately half a standard deviation below the mean established by their non-Hispanic, White American counterparts (Neisser et al., 1996; Puente & Salazar, 1998). Difficulty or strain in lexical decision making can be attributed to the amount of time spent using each language. As bilinguals use two languages, they may use each language less often than monolinguals who only use one language (Bialystok, 2009).

Statement of Purpose

The present study examines how bilingual Mexican Americans perform on neuropsychological measures of effort using optimal/suboptimal effort cutoffs that were derived from monolingual English speakers. It is important to identify measures that can appropriately diagnose bilingual patients as giving forth suboptimal effort. The TOMM and Dot Counting Test were chosen because they are two of the most widely used stand-alone SVTs that show high sensitivity (proper identification of suboptimal effort) and specificity (proper identification of optimal effort). In addition, the Reliable Digit Span is one of the most widely used embedded measures, again, because of the high level of sensitivity and specificity that it demonstrates. The language format and characteristics of a neuropsychological test can affect how well bilinguals perform. Similarly, there are visual-perceptual assessments that tap into executive functioning whereby bilingual participants must distinguish between conflicting stimuli that should present a problem in processing the correct responses as a result of interference. However, bilinguals actually score comparably, if not better than monolinguals on said tasks (Paap, Johnson, & Sawi, 2015). This study will further add to the growing body of knowledge regarding neuropsychological testing on linguistically diverse populations and what types of assessments work best on them. The following hypotheses will be tested in order to carry out the purpose of this study.

Hypotheses

Hypothesis 1

Mexican American bilinguals will produce more false positives (misidentification of suboptimal effort), as derived from optimal/suboptimal effort cut-offs developed using monolingual English speakers, on a language-format embedded test of effort, Reliable Digit Span.

Hypothesis 2

Mexican American bilinguals will produce similar numbers of false positives (misidentification of suboptimal effort), as derived from optimal/suboptimal effort cut-offs developed using monolingual English speakers, on the visual-perceptual format tests of effort, TOMM and the Dot Counting Test.

Hypothesis 3

Age will have a negative relationship with education and performance on the following measures of effort: TOMM and Reliable Digit Span. However, due to scoring criteria, there will be a positive relationship between age and scores on the Dot Counting Test. Additionally, there will be positive relationships between demographic variables (education and income) and performance on two measures of effort (TOMM, Reliable Digit Span) Due to the scoring criteria, there will be a negative relationship between demographic variables (education and income) and the Dot Counting Test. Relationships between demographic variables and measures of effort will be explored.

CHAPTER III

METHODOLOGY

Method

Participants

Consecutive participants (*N* = 66) were recruited by word-of-mouth from the Rio Grande Valley. They were subjectively bilingual Mexican Americans, residents of the U.S., aged $\geq$18 years, with $\leq$ 15 years of education, not currently enrolled in college, and without history of neurological disorders (e.g., stroke), psychiatric disorders (e.g., depression), or drug/alcohol abuse.

Of the 66 participants, 36 (55%) were male and 30 females (45%). The age range was 18 to 89 years (*M* = 39.09; *SD*=15.97). There was a total of 29 (44%) participants born in Mexico, while 37 (56%) had been born in the U.S. Those born in Mexico had resided in the U.S. from 1 to 59 years (*M* = 26.86; *SD* = 14.56). Participants' level of education ranged from 4 to 15 years (*M* = 11.44; *SD* = 1.91). There were 44 participants (66.7%) educated in the U.S., 17 (25.8%) in Mexico, and 5 (7.6%) in both countries. About half of the participants reported that their yearly household income was less than $30,000, whereas the median household income in Hidalgo County currently stands as $33,218. A majority of the participants, 55 (83%), learned Spanish as their first language, while 8 (12%) learned English first, and 3 (5%) reported to have learned both languages simultaneously. Age of second language acquisition ranged from 2 to 29 years

(*M* = 8.26; *SD* = 6.74). Currently, 46% of participants preferred speaking both languages, 30% preferred Spanish, and 24% English. Language spoken at homes was predominantly Spanish (62%), followed by English (33%). The remainder (5%) spoke both Spanish and English at home. There were 35 (53%) participants that preferred to be tested in English and 31 (47%) that preferred Spanish.

Measures

Demographic form. A demographics survey requested information on: gender; date of birth; age; country of birth; number of years in the U.S.; number of years of education; country of education; occupation; household income; preferred language for conversation; what language was learned first as a child; when second language acquisition began; language predominantly spoken at home; and history of neurological disorder, psychiatric disorder or drug/alcohol abuse.

Woodcock-Munoz Language Survey-Revised. The Picture Vocabulary subtest of the Woodcock-Munoz Language Survey-Revised (WMLS-R: Woodcock et al., 2005) in both English and Spanish was administered to provide a measure of language fluency and bilingualism. The subtest was scored according to standardized procedures wherein raw scores were converted into standardized scores (derived from a national sample composed of English- and Spanish-speaking monolinguals: *M* = 100; *SD* = 15).

In order to provide a measure of bilingualism, the standardized Picture Vocabulary subtest score of the WMLS-R in English was subtracted from its Spanish counterpart. Participants were classified into groupings according to their degree of bilingualism. Balanced bilinguals scores ranged from -10 to 10, while those who were Spanish-dominant bilingual scored > 10, whereas English-dominant bilinguals scored < -10. There were 13 (20%) participants who were English-dominant bilingual, 12 (18%) Spanish-dominant bilingual, and 41 (62%) balanced bilingual. Of the balanced bilinguals, 21 (51%) were tested in English and 20 (49%) were tested in Spanish. There were 25 language-dominant bilinguals (combining both English and Spanish-dominant), of those 14 (56%) were tested in English and 11 (44%) in Spanish.

Independent samples *t*-tests were conducted to examine if there was a difference in participants' age, education, and income among the following groups: (a) English vs. Spanish language of administration; (b) balanced bilingual vs. language-dominant bilingual when tested in English; and (c) balanced bilingual vs. language-dominant bilingual when tested in Spanish.

Participants who were tested in English were 36.49 years old (*SD* = 12.60) while those tested in Spanish were 42.03 years old (*SD* = 18.87). There was no significant difference in age: t (64) = -1.42. p = .161. Participants that were tested in English had 11.83 years of education (*SD*= 1.01) and a yearly income between $20,000 and $30,000, while those tested in Spanish had 11 years of education (*SD*= 2.52) and a yearly income between $30,000 and $40,000. There was no significant difference in education: t (64) = 1.79, p = .078; or income: t (62) = -1.12, p = .27.

Balanced bilinguals who were tested in English were 41.10 years old (*SD* = 12.49), with 11.67 years of education (*SD* = 1.02), and between $30,000 and $40,000 of yearly income while language dominant bilinguals tested in English were 29.57 years old (*SD*= 9.46) with 11.67 years of education (*SD*= .99) and a yearly income between $10,000 and $20,000. There was a significant difference in age between balanced bilinguals and language dominant bilinguals when tested in English, wherein balanced bilinguals were older than language dominant bilinguals: t (33) = -2.93, p = .006. There was no significant difference in education: t (33) = 1.16, p = .25; or income: t (32) = -1.61, p = .12, between balanced bilinguals and language dominant bilinguals when tested in English.

Balanced bilinguals who were tested in Spanish were 37.95 years old (*SD* = 14.41), with 11.47 years of education (*SD* = 2.32) and between $40,000 and $50,000 of yearly income while language dominant bilinguals tested in Spanish were 48.50 years old (*SD*= 23.59) with 10.25 years of education (*SD = 2*.73) and a yearly income between $20,000 and $30,000. There was no significant difference between balanced bilinguals and language dominant bilinguals in age: $t(29) = 1.39$, *ns*; education: $t(29) = -1.52$, ns; and income: $t(28) = -1.37$, *ns*, when tested in Spanish.

Test of Memory Malingering (TOMM). The Test of Memory Malingering (TOMM: Tombaugh, 1996) is a 50-item, forced-choice, visual-perceptual format, memory test that assesses participant effort. Participants are shown a series of 50 pictures for three seconds each that they must later recognize between two choices on the subsequent recognition trial. There are two learning trials, each followed by a recognition trial, and an optional delayed retention trial. The recommended criterion cutoff within the TOMM manual in order to detect suboptimal effort is either: (a) $\leq$ 90% correct on the second trial or the retention trial; or (b) a score lower than chance performance (i.e., < 50%) on any of the three trials (Tombaugh, 1996).

Reliable Digit Span. Reliable Digit Span is a language format task that measures the maximum number of digits that were correctly repeated over two trials each both forwards and backwards. In order to detect suboptimal effort, a cutoff score of $\leq$ 7 is recommended (Greiffenstein et al., 1994). Using this cutoff score, Meyers and Volbrecht (1998) found that Reliable Digit Span classified 49% of litigants, but only 4% of non-litigants referred for neuropsychological assessment.

Dot Counting Test. The Dot Counting Test is a visual-perceptual format task that consists of 12 cards of grouped and ungrouped dots that participants must count as quickly as they can. E-scores for this assessment are derived from the combination of average time for grouped and ungrouped dots in seconds plus the total number of errors. The testing manual recommends the use of a combination score of $\geq$ 14 to define suboptimal effort. This cutoff score gave a false positive rate of 12% in a nonclinical, monolingual, English-speaking sample (Boone et al., 2002). This proposed method of scoring has been found more reliable in recent literature as opposed to the traditional method wherein signs of suboptimal effort were detected if the amount of time needed to count the grouped dots surpassed the amount of time needed to count the dispersed ones (Vilar-Lopéz et al., 2008).

Procedure. At the beginning of every session participants were given consent forms in either English or Spanish. Bilingual research assistants began with the Picture Vocabulary subtest of the WMLS-R in either English or Spanish (they alternated the starting language for every other participant). Assessments were continued in the participant's preferred language. The first two trials of the TOMM were then administered, followed by Reliable Digit Span. Thereafter participants completed the Dot Counting Test and the retention trial of the TOMM. In total, the entire session lasted between 30 and 40 minutes

CHAPTER IV

RESULTS

Means (*SD*s) for the following variables are reported in Table 1: TOMM Trial 2; Reliable Digit Span; Dot Counting; WMLS-R English Score; WMLS-R Spanish Score; and the absolute value of degree of bilingualism. Using the recommended criterion cutoff within the TOMM manual to detect suboptimal effort: (a) ≤ 90% correct on the second trial or the retention trial; and (b) a score lower than chance performance (i.e., < 50%) on any of the three trials, 5% of bilingual Mexican American scores were false positives. Reliable Digit Span had 36% of the scores result in false positives when using the recommended optimal/suboptimal effort score ≤ 7. Using an optimal/suboptimal effort cutoff E-score of ≥14, the Dot Counting Test resulted in 12% false positives. The percentage of false positives that are derived from the TOMM manual, using monolingual English-speakers, is 1% (Tombaugh, 1996). The Dot Counting Test manual had 12% of its scores, derived from monolingual English- speakers, be false positives (Boone et al., 2002). The Reliable Digit Span had no false positives when testing a non-clinical, monolingual English-speaking, sample of 50 undergraduate psychology students (Silk-Eglit et al., 2014).

Table 1

Means (SDs) for the Test of Memory Malingering (TOMM) Trial 2, Reliable Digit Span, Dot Counting, Woodcock Munoz Language Survey-Revised (WMLS-R) English Score, WMLS-R Spanish score, and degree of bilingualism

	N	*M*	*SD*
TOMM Trial 2	66	48.67	4.74
Reliable Digit Span	66	8.52	2.02
Dot Counting	66	10.44	3.51
WMLS-R English Score	66	83.85	12.01
WMLS-R Spanish Score	66	85.09	6.94
Degree of Bilingualism	66	10.52	10.47

Hypothesis Testing

Hypothesis 1

It was hypothesized that bilingual Mexican Americans would produce more false positives, as derived from optimal/suboptimal effort cut-offs developed using monolingual English speakers, on the Reliable Digit Span. Using the recommended criterion cutoff of ≤ 7, this sample of bilingual Mexican Americans had 36% of scores result in false positives. For the sake of comparison, Silk-Eglit et al. (2014) had no false positives in their sample.

As a way to further explore the false positives on the Reliable Digit Span, Chi Squares were conducted to examine if there were significant difference in the proportion of false positives between: (a) English vs. Spanish language of administration (see Table 2); (b) balanced vs.

language-dominant bilinguals when tested in English (see Table 3); and (c) balanced vs. language-dominant bilinguals when tested in Spanish (see Table 4).

Table 2

Total Number of False Positives on the Reliable Digit Span between English vs. Spanish Language of Administration

	English	Spanish	Total
Optimal effort	24	18	42
Suboptimal effort (false positive)	11	13	24
			66

There was no difference in the proportion of false positives in English (31%) vs. Spanish (42%) language of administration: χ^2 (1, N = 66) = .78, p = .27

Table 3

False Positives on the Reliable Digit Span between Balanced and Language Dominant Bilinguals Tested in English

	Balanced	Language Dominant	Total
Optimal effort	12	12	24
Suboptimal effort (false positive)	9	2	11
			35

For participants that were tested in English, there was no difference in the percentage of false positives for balanced bilingual (43%) and language-dominant bilinguals (14%): χ^2 (1, N = 66) = 3.18, p = .07.

Table 4
False Positives on the Reliable Digit Span between Balanced and Language-Dominant Bilinguals Tested in Spanish

	Balanced	Language Dominant	Total
Optimal effort	7	11	18
Suboptimal effort (false positive)	5	8	13
			31

There was no difference in proportion of false positives between balanced bilinguals (42%) and language-dominant bilinguals (43%) when tested in Spanish: χ^2 (1, N = 66) = .0001. p = .64.

Hypothesis 2

It was hypothesized that bilingual Mexican Americans would produce similar number of false positives on the TOMM and the Dot Counting Test, as derived from optimal/suboptimal effort cut-offs developed using monolingual English-speakers Using the recommended criterion cutoff within the TOMM manual in order to detect suboptimal effort: (a) $\leq$ 90% correct on the second trial or the retention trial; and (b) a score lower than chance performance (i.e., < 50%) on any of the three trials, 5% of bilingual Mexican American scores were false positives.

Using an E-Score of $\geq$ 14 on the Dot Counting Test, 12% of scores were false positives. The respective manual for the TOMM (Tombaugh, 1996) and the Dot Counting Test (Boone et al., 2002) had a 1% and 12% rate of false positives from samples of monolingual, English-speakers.

Hypothesis 3

Pearson's product-moment correlational analyses were conducted to examine the interrelations among demographic variables (age, years of education, income), scores on measures of effort (TOMM Trial 2, Reliable Digit Span, Dot Counting Test), WMLS-R Picture Vocabulary (Spanish and English), and degree of bilingualism (Table 5). When variables were found to significantly correlate with the criterion variable (Reliable Digit Span), they were included as independent predictors in a linear regression to determine how each predictor(s) uniquely and collectively accounted for variance in the likelihood of producing low scores on Reliable Digit Span.

Table 5

Intercorrelations among Age, Education, Income, Test Of Memory Malingering (TOMM) Trial 2, Reliable Digit Span, Dot Counting, Woodcock Munoz Language Survey-Revised (WMLS-R) English Score, WMLS-R Spanish Score, and degree of Bilingualism

	1	2	3	4	5	6	7	8	9
1.Age	-	-.57**	.03	-.31*	-.38**	.31*	.06	.31*	-.21
2.Education		-	.04	.55**	.13	-.33**	.13	-.21	-.05
3.Income			-	-.01	-.22	-.18	.09	.24	-.24
4.TOMM Trial 2				-	.15	-.51**	-.03	.02	.09
5.Reliable Digit Span					-	-.28*	.15	-.31*	.07
6.Dot Counting						-	.01	.00	-.02
7.WMLS-R English							-	-.17	-.59**
8.WMLS-R Spanish								-	-.10
9. Degree of Bilingualism									-

* $p < .05$
** $p < .01$

It was hypothesized that age would have a negative relationship with education and performance on TOMM and Reliable Digit Span. Due to the scoring criteria, age would have a positive relationship with the Dot Counting Test. Furthermore, there would be positive relationships between demographic variables (education and income) and performance on TOMM and Reliable Digit Span. Again, due to the scoring criteria, there would be a negative relationship between demographic variables (education and income) and the Dot Counting Test.

There were negative relationships between age and the following variables: (a) education, r (64) = -.57, $p < .01$; (b) TOMM Trial 2, r (64) = -.31, $p < .05$; and (c) Reliable Digit Span, r (64) = -.38, $p < .01$, where older participants had less education and lower scores on the measures of effort. A positive relationship was found between age and scores on the Dot Counting Test, r (64) = .31, $p < .05$, where older participants had higher scores. Additionally, there was a positive relationship between age and WMLS-R Spanish score, r (64) = .31, $p < .05$, where older participants had higher Spanish language scores.

Education was positively correlated with TOMM Trial 2, r (64) = .55, $p < .01$, where participants with more education had higher scores on TOMM Trial 2. A negative relationship between education and the Dot Counting Test was found, r (64) = -.33, $p < .01$, where participants with more education had lower scores on Dot Counting. A negative correlation was found between scores on TOMM Trial 2 and their scores on the Dot Counting Test, r (64) = -.51, $p < .01$, meaning that higher scores on TOMM Trial 2 would lead to lower scores on Dot Counting.

The Reliable Digit Span had a negative relationship between both the Dot Counting Test, r (64) = -.28, $p < .05$ and the Spanish WMLS-R score r (64) = -.31, $p < .05$. Higher scores on the Reliable Digit Span lead to lower Dot Counting Test scores and lower WMLS-R Spanish score. Lastly, there was a negative relationship between WMLS-R English score and the participants' degree of bilingualism, wherein the better their score in English was the more likely that they a balanced bilingual.

A multiple linear regression analysis with participant age, Dot Counting Test scores, and WMLS-R Spanish scores as predictors of scores on Reliable Digit Span was conducted. Collectively the predictors accounted for a significant proportion of the variance in Reliable Digit Span, $R^2 = .23$, $F\ (3, 62) = 6.08$, $p = .001$. Age was not a predictor of low scores on Reliable Digit Span, $\beta = -.24$, $t\ (62) = -1.94$, $p = .55$. Scores on the Dot Counting Test were also not predictors of low scores on Reliable Digit Span: $\beta = -.21$, $t\ (62) = -1.76$, $p = .084$. WMLS-R Spanish Scores were inversely related to scores on the Reliable Digit Span: $\beta = -.24$, $t\ (62) = -2.03$, $p < .05$.

CHAPTER V

DISCUSSION

The present study examined how neurocognitively intact, bilingual, Mexican Americans performed on three of the most widely used neuropsychological measures of effort: TOMM; Reliable Digit Span; and Dot Counting. Participant effort was measured using optimal/suboptimal effort cut-offs developed using monolingual English-speakers. The first two hypotheses addressed the question of whether these measures were appropriate for bilingual Mexican Americans with regards to the proper detection of their effort level. The first hypothesis focused on a language format embedded measure of effort, the Reliable Digit Span. Because it is a language format assessment that measures the maximum amount of digits one can recall both forwards and backwards, it was expected that bilingual Mexican Americans would produce a higher number of false positives as compared to monolingual English-speakers. The second hypothesis focused on two visual-perceptual format, stand-alone measures of effort, the TOMM and Dot Counting Test. It was expected that bilingual Mexican Americans would produce false positives at similar rates on TOMM and Dot Counting when compared to monolingual English-speakers because they use visual-perceptual formats.

Results supporting the first hypothesis showed that this bilingual Mexican American sample did indeed produce a higher rate of false positives (36%) when using the suggested criterion cutoff score of ≤ 7, on the language format embedded measure of Reliable Digit Span as compared to monolingual English-speaking undergraduate students (0%: Silk-Eglit et al., 2014).

To further explore the rate of false positives on the Reliable Digit Span, three comparisons were made: (a) English vs. Spanish language of administration; (b) balanced bilingual vs. language-dominant bilingual when tested in English; and (c) balanced bilingual vs. language-dominant bilingual when tested in Spanish. These comparison groups were not significantly different in age, education, and income, except that the balanced bilingual group was older than the language-dominant bilinguals when tested in English. There were no differences in the proportion of false positives among any of these three comparisons, meaning that the false positives that were produced on the Reliable Digit Span did not differ significantly by language of administration or bilingual grouping.

Reliable Digit Span is a language format measure of effort. The rate of false positives on this assessment with a bilingual, Mexican American sample (36%) suggests that its usage with this linguistically diverse group leads to greater misdiagnoses of suboptimal effort when compared to the monolingual English-speakers. This may be because it taxes processing capacity limits within the language domain. Reliable Digit Span asks participants to recall an increasingly longer list of numbers that was read aloud to them both forwards and backwards. Competition between languages, because of the simultaneous processing of two languages may cause poor performance on language assessments (Mind et al., 2008). Likewise, the amount of time an individual spends in either language can cause poor performance because of the difference in language proficiency that is needed for advanced cognitive and linguistic skills (Mind et. al., 2008). Optimal/suboptimal effort cutoff scores on Reliable Digit Span may not be appropriate for bilingual Mexican Americans because they do not represent the majority population on which the optimal/suboptimal effort cutoffs were established (Elbulok-Charcape et al., 2014). Results supporting the second hypothesis showed that this bilingual Mexican American sample produced similar percentages of false positives on the visual-perceptual format stand-alone measures of effort: TOMM and Dot Counting, using the recommended cutoff scores derived from monolingual English-speakers. Using the recommended criterion cutoff within the TOMM manual in order to detect suboptimal effort: (a) ≤ 90% correct on the second trial or the retention trial; and (b) a score lower than chance performance (i.e., < 50%) on any of the three trials, 5% of bilingual Mexican American scores were false positives, whereas monolingual English-speakers had a 1% rate.

In addition, when using an E-score of ≥ 14 on the Dot Counting test, 12% of participants' scores were false positives as compared to 12% with monolingual English-speakers. Results suggest that the usage of these stand-alone measures of effort are indeed accurate and appropriate for bilingual Mexican Americans. Findings from the present study further corroborate results reported by Vilar-Lopéz et al. (2008), who examined the usage of TOMM and Dot Counting to detect malingering in a Spanish-speaking population from Spain. Both assessments were able to differentiate between the groups of malingerers and non-malingerers. The present study extends past results and suggests that the TOMM is an effective effort assessment tool for use with English- and Spanish-speaking bilinguals.

Hypothesis 3 examined relations among demographic variables (age, education, income), measures of effort (TOMM Trial 2, Reliable Digit Span, Dot Counting Test), WMLS-R English and Spanish score, and bilingualism. It was hypothesized that age would have a negative relationship with education and performance on TOMM and Reliable Digit Span. Due to the scoring criteria, age would have a positive relationship with the Dot Counting Test. Furthermore, there would be positive relationships between demographic variables (education and income) and performance on TOMM and Reliable Digit Span. Again, due to the scoring criteria, there would be a negative relationship between demographic variables (education and income) and the Dot Counting Test. Relationships among demographic variables, measures of effort, and bilingualism were to be explored.

Age was inversely related to the following variables: (a) education; (b) TOMM Trial 2; (c) and Reliable Digit Span. Older participants were more likely to have less education and have lower scores on TOMM Trial 2 and Reliable Digit Span. However, older participants had higher Dot Counting Test scores as higher scores indicate poorer performance. Lower scores on TOMM Trial 2 and Reliable Digit Span and higher scores on the Dot Counting Test are indicative of poor performance which can be misinterpreted as suboptimal performance. Additionally, there was a positive relationship between age and WMLS-R scores in Spanish, where older participants had better Spanish scores that indicates better proficiency in the Spanish language. Thus, participants' age and education should be considered when using tests of effort and language proficiency.

Participants with more education had higher scores on TOMM Trial 2. Conversely though, participants with more education had lower Dot Counting Test scores. Because of the inverse meaning of test scores on TOMM Trial 2 and Dot Counting test, participants with more education had better performance which indicates optimal effort. Participants with higher TOMM Trial 2 scores had lower Dot Counting Test scores, both being representative of better performance. It was interesting to note that income did not correlate with any other variable whereas education did correlate with TOMM Trial 2 and Dot Counting. Both variables are part of socio-economic status and as a result, warrant further investigation as to why there were no correlations with income.

Higher scores, which are indicative of optimal effort, on Reliable Digit Span led to lower Dot Counting Test scores (better performance) and lower WMLS-R Spanish scores. Thus, better performance on Reliable Digit Span was associated with better performance on the Dot Counting test, but a lower degree of Spanish language proficiency. Inversely though, lower scores on Reliable Digit Span would be indicative of a higher degree of Spanish language proficiency.

Reliable Digit Span, a language format assessment, yielded a false positive rate of 36% (misidentification of suboptimal effort) with this sample of bilingual Mexican Americans. It is important to further investigate the relationship that the language format embedded measure, Reliable Digit Span, may have on acculturative and other bilingual variables. These relationships can lead to further understanding the potential impact factors on optimal/suboptimal effort measures for linguistically diverse populations.

Greater Spanish language proficiency predicted lower scores on the Reliable Digit Span. This further supports the results from hypothesis 1 that suggest that the Reliable Digit Span is not a proper assessment to use on bilingual Mexican Americans because of the high rate of misidentification of this particular ethnic group as giving suboptimal effort. It is important to point out that age was very close to being a significant independent predictor of lower scores on Reliable Digit Span. Thus future research should include a larger sample size to increase statistical power and examine the relationship between age and performance on the Reliable Digit Span test.

The TOMM and the Dot Counting Test show comparable results for bilinguals and monolingual English-speakers. Whereas the Reliable Digit Span, because of its language format and subsequent effects on bilinguals, produced a high rate of false positives (36%) on bilingual Mexican Americans when compared to monolingual English speakers (0%). Different approaches to examining language and its effects on neuropsychological testing and Symptom Validity Testing should be explored to increase our understanding of the effects of bilingualism on neurocognitive and effort tests. Findings from more studies on the reliability and validity of assessments can further be used to help patients receive proper evaluation.

Future studies should also investigate the effects of acculturation on demographic variables and scores on neurocognitive and effort assessments. The different levels of acculturation, along with degree of bilingualism, and language proficiency can affect results on cognitive tests for Hispanics (Gasquoine, 1999). In addition, increasing the sample size for the different types of bilingual groupings such as: (a) balanced bilingual; (b) English-dominant bilingual; and (c) Spanish-dominant bilingual, would further increase the amount of variance between and within the groups. As a result, better comparisons can be made among the bilingual groupings. This study included balanced bilinguals and language-dominant bilinguals, with the latter being the combination of English- and Spanish-dominant bilinguals. Recreating this study with different measures of effort is also encouraged to explore how other tests of effort are appropriate for use with bilingual Mexican Americans. More specifically, an equal amount of visual-perceptual and language format measures should be used to further corroborate the findings. Ultimately, this will contribute to the practical usefulness of Symptom Validity Testing with Spanish-English bilinguals.

REFERENCES

Armistead-Jehle, P., & Hansen, C. L. (2011). Comparison of the Repeatable Battery for the assessment of Neuropsychological Status Effort Index and stand-alone symptom validity tests in a military sample. *Archives of Clinical Neuropsychology*, *26*(7), 592-601.

Arnold, B. R., Montgomery, G. T., Castañeda, I., & Longoria, R. (1994). Acculturation and performance of Hispanics on selected Halstead-Reitan neuropsychological tests. *Assessment*, *1*(3), 239-248.

Axelrod, B. N., Fichtenberg, N. L., Millis, S. R., & Wertheimer, J. C. (2006). Detecting Incomplete Effort with Digit Span from the Wechsler Adult Intelligence Scale—Third Edition. *The Clinical Neuropsychologist*, *20*(3), 513-523.

Bialystok, E. (1999). Cognitive complexity and attentional control in the bilingual mind. *Child Development*, *70*(3), 636-644.

Bialystok, E., Luk, G., & Kwan, E. (2005). Bilingualism, biliteracy, and learning to read: Interactions among languages and writing systems. *Scientific Studies of Reading*, *9*(1), 43-61.

Bialystok, E. (2009). Bilingualism: The good, the bad, and the indifferent. *Bilingualism: Language and cognition*, *12*(1), 3-11.

Bianchini, K. J., Mathias, C. W., & Greve, K. W. (2001). Symptom validity testing: A critical review. *The Clinical Neuropsychologist*, *15*(1), 19-45.

Boone, K. B., Lu, P., Sherman, D., Palmer, B., Back, C., Shamieh, E., & Berman, N. G. (2000). Validation of a new technique to detect malingering of cognitive symptoms: The b Test. *Archives of Clinical Neuropsychology*, *15*(3), 227-241.

Boone, K., Lu, P., & Herzberg, D. (2002). *The Dot Counting Test*. Los Angeles: Western Psychological services.

Boone, K. B., Lu, P., Back, C., King, C., Lee, A., Philpott, L., … Warner-Chacon, K. (2002). Sensitivity and specificity of the Rey Dot Counting Test in patients with suspect effort and various clinical samples. *Archives of Clinical Neuropsychology*, *17*(7), 625-642.

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.

Conners, C. K., Staff, M. H., Connelly, V., Campbell, S., MacLean, M., & Barnes, J. (2000). *Conners' continuous performance Test II (CPT II v. 5)*. Toronto, ON: Multi-Health Systems Inc.

Delis, D. C., Kramer, J. H., Kaplan, E., & Ober, B. A. (2000). *California Verbal Learning Test-II*. NY: Psychological Corporation.

Elbulok-Charcape, M. M., Rabin, L. A., Spadaccini, A. T., & Barr, W. B. (2014). Trends in the neuropsychological assessment of ethnic/racial minorities: A survey of clinical neuropsychologists in the United States and Canada. *Cultural Diversity and Ethnic Minority Psychology*, *20*(3), 353-361.

Gasquoine, P. G. (1999). Variables moderating cultural and ethnic differences in neuropsychological assessment: The case of Hispanic Americans. *The Clinical Neuropsychologist*, *13*(3), 376-383.

Gasquoine, P. G. (2013). *Neuropsychology: Cellular physiology to clinical practice* (2nd ed.). Redding, CA: BVT Publishing.

Gasquoine, P. G., Croyle, K. L., Cavazos-Gonzalez, C., & Sandoval, O. (2007). Language of administration and neuropsychological test performance in neurologically intact Hispanic American bilingual adults. *Archives of Clinical Neuropsychology*, *22*(8), 991-1001.

Gasquoine, P. G., & Gonzalez, C. D. (2012). Using monolingual neuropsychological test norms with bilingual Hispanic Americans: Application of an individual comparison standard. *Archives of Clinical Neuropsychology, 27,* 268-276.

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and cognition*, *1*(2), 67-81.

Greiffenstein, M. F., Baker, W. J., & Gola, T. (1994). Validation of malingered amnesia measures with a large clinical sample. *Psychological Assessment, 6*, 218-224.

Heaton, R. K., Chelune, G. J., Talley, J. L., Kay, G. G., & Curtiss, G. (1993). Wisconsin card sort test manual: Revised and expanded. *Odessa*, *FL: Psychological Assessment Resources.*

Heilbronner, R. L., Sweet, J. J., Morgan, J. E., Larrabee, G. J., Millis, S. R., & Conference Participants, (2009). American Academy of Clinical Neuropsychology Consensus Conference Statement on the neuropsychological assessment of effort, response bias, and malingering. *The Clinical Neuropsychologist*, *23*(7), 1093-1129.

Inman, T. H., & Berry, D. T. (2002). Cross-validation of indicators of malingering: A comparison of nine neuropsychological tests, four tests of malingering, and behavioral observations. *Archives of Clinical Neuropsychology*, *17*(1), 1-23.

Larrabee, G. J. (1992). Interpretive strategies for evaluation of neuropsychological data in legal settings. *Forensic Reports*, 5(3), 257-264.

Larrabee, G. J. (2003). Detection of malingering using atypical performance patterns on standard neuropsychological tests. *The Clinical Neuropsychologist*, *17*(3), 410-425.

Loughan, A. R., Perna, R., & Hertza, J. (2012). The value of the Wechsler intelligence scale for children-digit span as an embedded measure of effort: An investigation into children with dual diagnoses. *Archives of Clinical Neuropsychology*, *27*(7), 716-724.

McMillan, T. M., Anderson, S., Baker, G., Berger, M., Powell, G. E., & Knight, R. (2009). Assessment of effort in clinical testing of cognitive functioning for adults. *Professional Practice Board-British Psychological Society*, *1*, 1-27.

Meyers, J. E., & Volbrecht, M. (1998). Validation of reliable digits for detection of malingering. *Assessment*, *5*(3), 303-307.

Meyers, J. E., Volbrecht, M., Axelrod, B. N., & Reinsch-Boothby, L. (2011). Embedded symptom validity tests and overall neuropsychological test performance. *Archives of Clinical Neuropsychology*, *26*(1), 8-15.

Mindt, M. R., Arentoft, A., Germano, K. K., D'Aquila, E., Scheiner, D., Pizzirusso, M., & Gollan, T. H. (2008). Neuropsychological, cognitive, and theoretical considerations for evaluation of bilingual individuals. *Neuropsychology Review*, *18*(3), 255-268.

Mungas, D., Reed, B. R., Haan, M. N., & González, H. (2005). Spanish and English neuropsychological assessment scales: relationship to demographics, language, cognition, and independent function. *Neuropsychology*, *19*(4), 466.

Neisser, U., Boodoo, G., Bouchard Jr, T. J., Boykin, A. W., Brody, N., Ceci, S. J.,...Urbina, S. (1996). Intelligence: knowns and unknowns. *American Psychologist*, *51*(2), 77-101.

Novitski, J., Steele, S., Karantzoulis, S., & Randolph, C. (2012). The repeatable battery for the assessment of neuropsychological status effort scale. *Archives of Clinical Neuropsychology*, *27*(2), 190-195.

Paap, K. R., Johnson, H. A., & Sawi, O. (2015). Bilingual advantages in executive functioning either do not exist or are restricted to very specific and undetermined circumstances. *Cortex*, *69*, 265-278.

Paulson, D., Horner, M. D., & Bachman, D. (2015). A comparison of four embedded validity indices for the RBANS in a memory disorders Clinic. *Archives of Clinical*

Neuropsychology, *30*(3), 207-216.

Pérez-Arce, P. (1999). The influence of culture on cognition. *Archives of Clinical Neuropsychology*, *14*(7), 581-592.

Pontón, M. O., & Ardila, A. (1999). The future of neuropsychology with Hispanic populations in the United States. *Archives of Clinical Neuropsychology*, *14*(7), 565-580.

Puente, A. E., & Salazar, G. D. (1998). Assessment of Minority and Culturally Diverse Children. In A. Prifitera & D. H. Saklofske (Eds.), *WISC-III Clinical Use and Interpretation* (1st ed., pp. 227-248). San Diego, CA: Academic Press.

Puente, A. E., & Perez-Garcia, M. (2000). Psychological assessment of ethnic minorities. In G. Goldstein & M. Hersen (Eds.), *Handbook of Psychological Assessment* (3rd ed.). Chicago: Pergamon Press, pp. 527-551.

Rey, A. (1941). L'examen psychologique dans les cas d'encéphalopathie traumatique. (Les problems.). *Archives de Psychologie*, *28*(112), 286–340.

Rey, A. (1964). *L'examen clinique en psychologie*. Paris: Presses universitaires de France.

Silk-Eglit, G. M., Stenclik, J. H., Gavett, B. E., Adams, J. W., Lynch, J. K., & McCaffrey, R. J. (2014). Base rate of performance invalidity among non-clinical undergraduate research participants. *Archives of Clinical Neuropsychology*, *29*(5), 415-421.

Slick, D. J., Sherman, E. M., & Iverson, G. L. (1999). Diagnostic criteria for malingered neurocognitive dysfunction: Proposed standards for clinical practice and research. *The Clinical Neuropsychologist*, *13*(4), 545-561.

Tombaugh, T. N. (1996). *Test of Memory Malingering.* North Tonawanda, NY: Multi-Health Systems.

Trahan, D. E., & Larrabee, G. J. (1988). *Continuous visual memory test: Professional manual*. Odessa, FL: Psychological Assessment Resources.

Vickery, C. D., Berry, D. T., Inman, T. H., Harris, M. J., & Orey, S. A. (2001). Detection of inadequate effort on neuropsychological testing: A meta-analytic review of selected procedures. *Archives of Clinical Neuropsychology*, *16*(1), 45-73

Vilar-López, R., Gómez-Río, Caracuel-Romero, A., Llamas-Elvira, J., & Perez-Garcia, M. (2008). Use of specific malingering measures in a Spanish sample. *Journal of Clinical and Experimental Neuropsychology*, *30*(6), 710-722.

Vilar-López, R., Gómez-Río, M., Santiago-Ramajo, S., Rodríguez-Fernández, A., Puente, A.E., & Pérez-García, M. (2008). Malingering detection in a Spanish population with a known-groups design. *Archives of Clinical Neuropsychology*, *23,* 365-77.

Vilar-López, R., & Puente, A. E. (2010). Forensic neuropsychological assessment of members of minority groups: The case for assessing Hispanics. In A. MacNeill Horton & L. C. Hartlage (Eds.), *Handbook of Forensic Neuropsychology* (2nd ed.). New York, NY: Springer Publishing Company, pp. 309-332.

Wechsler, D. (1991). *WISC-III: Wechsler intelligence scale for children, 3rd edition: Manual*. NY: Psychological Corporation.

Wechsler, D. (2008). *WAIS-IV: Wechsler adult intelligence Scale: Technical and interpretive manual (4th ed.)*. San Antonio, TX: Pearson.

Woodcock, R. W., Munoz-Sandoval, A. F., Ruef, M. L., & Alvarado, C. G. (2005). *Woodcock-Munoz Language Survey-Revised*. Itasca, IL: Riverside Publishing.

BIOGRAPHICAL SKETCH

Arnoldo Amador was born in Edinburg, Texas. Arnoldo received his Bachelor of Arts in Psychology with a second major in English from the University of Texas - Pan American in August of 2013. He later earned his Master of Arts in Experimental Psychology from the University of Texas Rio Grande Valley in December of 2016. During his tenure as an undergraduate student, Arnoldo was involved in Sigma Tau Delta: English International Honor Society, serving as Historian and Vice President from its inception until his graduation in 2013. A little before becoming a graduate student, Arnoldo became involved in research and worked for various types of institutional grants such as: Title V, Howard Hughes Medical Institute (HHMI), and STEM Education funded by Department of Defense. In addition, Arnoldo was also part of many different research labs and has presented at various research conferences throughout the last two years. His permanent address is 3414 N. Jackson, Edinburg, TX 78541.