10-1-2021

Neurocognitive Profiles Associated with Limited English Proficiency in Cognitively Intact Adults

Maeme Adwoa Brantuo
University of Windsor

Neurocognitive Profiles Associated with Limited English Proficiency in Cognitively Intact Adults

By

Maame Adwoa Brantuo

A Thesis
Submitted to the Faculty of Graduate Studies
through the Department of Psychology
in Partial Fulfillment of the Requirements for
the Degree of Master of Arts
at the University of Windsor

Windsor, Ontario, Canada

2021

Neurocognitive Profiles Associated with Limited English Proficiency in Cognitively Intact Adults

by

Maame Adwoa Brantuo

APPROVED BY:

Z. Zhang
Faculty of Education

R. Biss
Department of Psychology

L. Erdodi, Advisor
Department of Psychology

July 12, 2021

DECLARATION OF ORIGINALITY

I hereby certify that I am the sole author of this thesis and that no part of this thesis has been published or submitted for publication.

I certify that, to the best of my knowledge, my thesis does not infringe upon anyone's copyright nor violate any proprietary rights and that any ideas, techniques, quotations, or any other material from the work of other people included in my thesis, published or otherwise, are fully acknowledged in accordance with the standard referencing practices. Furthermore, to the extent that I have included copyrighted material that surpasses the bounds of fair dealing within the meaning of the Canada Copyright Act, I certify that I have obtained a written permission from the copyright owner(s) to include such material(s) in my thesis and have included copies of such copyright clearances to my appendix.

I declare that this is a true copy of my thesis, including any final revisions, as approved by my thesis committee and the Graduate Studies office, and that this thesis has not been submitted for a higher degree to any other University or Institution.

ABSTRACT

A key assumption of neuropsychological testing in North America is that examinees are native speakers of English (NSE). The objective of the current study was to continue the line of investigations into the neurocognitive profiles associated with limited English proficiency (LEP). Participants were undergraduate students at a Canadian university. Data were collected from 40 NSEs and 40 participants with LEP. A battery of neuropsychological tests including measures with high (HVM) and low verbal mediation (LVM) was administered in counterbalanced order. As predicted, individuals with LEP performed more poorly on HVM measures and equivalent to NSEs on LVM measures, with some notable exceptions. Results suggest that clinicians should not interpret low scores on HVM tests as evidence of acquired cognitive impairment in individuals with LEP, since these measures may systematically underestimate cognitive ability in this population.

ACKNOWLEDGEMENTS

I would first like to thank my research advisor, Dr. Laszlo Erdodi, for his invaluable support and supervision over the last two years. His mentorship has helped keep me motivated and greatly increased my confidence in my abilities as a researcher. I would also like to thank my committee members, Dr. Renee Biss and Dr. Zuochen Zhang, for their thoughtful comments and contributions, which have helped refine this document into its current form.

Next, I would like to thank my cohort, who have been a source of tremendous support to me through the ups and downs of life in graduate school. I am incredibly blessed to have a group of such kind individuals in my life. A special thank you to the coven – Kat Matchett, Frasia Morrison, and Meredith Awrey – for their endless support, laughter, and willingness to proofread for me.

I would like to thank my parents, Yaw and (Dr.) Mary Brantuo, for their love, prayers, and all the sacrifices they have made to support my academic endeavors. Finally, to my sister, Afua Brantuo, thank you for being my biggest cheerleader and believing in me when I couldn’t believe in myself. I greatly appreciate you all, and this would not have been possible without you.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 1

Introduction

The Multilingual World

In a world that is becoming interconnected as a result of migration, technological advancements and the increasingly integrated global economy, it is no surprise that bilingualism is on the rise (Ryan, 2013). Based on statistics collected on bilingual populations in various regions of the world, it is estimated that more than half of the world's population is bilingual (Ansaldo et al., 2008). In Europe, for instance, 54% of the population reports speaking at least one additional language apart from their native tongue (Eurobarometer, 2012). This statistic increases to 80% when the analysis is restricted to the adults aged 25 to 64 (Eurostat, 2016).

The percentage of bilinguals is considerably smaller in North America in comparison to Europe but continues to grow at a steady pace. Crystal (1985) estimated that 235 million people around the world were English speaking bilinguals and that two thirds of children worldwide were being raised in bilingual environments. The American Community Survey reports that 20.14% of the US population is bilingual, a percentage that has almost doubled since 1980, mirroring the rapidly growing immigrant population (Ryan, 2013). In Canada, 35% of the population reports speaking more than one language, and a fifth reported having a mother tongue other than the official languages (i.e., English and French).

The Bilingual Brain: An Asset or a Liability?

The prevalence of bilingualism across the world inspired a considerable amount of research examining its mechanisms, cognitive benefits and disadvantages (Antoniou, 2019). An area of great debate appears to be the nature of the "bilingual advantage" in cognitive functioning, where there appears to be conflicting evidence on whether bilingualism enhances or interferes with cognitive processing. Existing empirical evidence is mixed. Some studies report a bilingual advantage (Adesope et al., 2010; Hilchey & Klein, 2011), but others have failed to replicate these findings (Paap & Greenberg, 2013), and not only question the concept of bilingual advantage (Samuel et al., 2018), but argue the opposite: bilingualism results in inefficiencies in certain cognitive processes (Ivanova & Costa, 2008; Paap et al., 2015). It has been hypothesized that this apparent bilingual advantage is due to the habitual recruitment of the executive control system (inhibition, planning, working memory, self-monitoring, switching) relative to monolinguals in order to resolve the conflict created by the joint activation of two languages (Bialystok, 2011). Consequently, the ongoing additional engagement of executive control for language switching eventually leads to improved executive function (Morrison et al., 2019).

Evidence of a Bilingual Advantage

There is a large body of evidence suggesting that bilingualism may be associated with positive cognitive outcomes. Adesope and colleagues (2010), for instance, conducted a systematic review and meta-analysis of 63 studies involving 6,022 participants. They found that bilingualism had a moderate positive effect (Hedges's g_s = 0.41) on various cognitive measures. Specifically, bilinguals demonstrated better performance than monolinguals on measures of attentional control, problem solving, metalinguistic skills, metacognitive awareness, abstract/symbolic representation skills as well as working memory.

Their results were replicated by Hilchey and Klein (2011), who found that bilinguals demonstrated an executive processing advantage on interference tasks such as the Simon task. However, the bilingual advantage was evident only in older adults, and not in other age groups. This is consistent with evidence that shows that bilingualism is associated with positive cognitive outcomes in aging. Bialystok et al. (2007) found that bilingual individuals tend to be diagnosed with Alzheimer's disease 4 years later than monolingual individuals. The delay in the onset of different forms of dementia in bilinguals is a reliable finding that has been replicated in different populations (Freedman et al., 2014; Woumans et al., 2015). Furthermore, this cognitive benefit in aging remains even in the absence of confounds like immigration status and level of education, suggesting that it is indeed due to bilingualism, and not contextual variables (Alladi et al., 2013).

Grundy and Timmer (2016) also reported a significant small to medium sized effect of bilingualism on working memory capacity. Additionally, they found that the language in which the task was performed moderated the relationship such that the effect size was smaller when the task was conducted in the participant's non-dominant language. Taken together, these findings support the notion of the bilingual advantage in cognitive performance, while also delineating some of its limits.

The Controversy around the Bilingual Advantage

Other researchers have failed to replicate the findings supporting the presence of a bilingual advantage in processing. Some particularly vocal opponents of the bilingual advantage included Paap and colleagues (2015) who argued that the majority of studies they examined failed to show convincing evidence for a bilingual advantage. Additionally, the studies that demonstrated differences had smaller sample sizes than the ones with negative findings. They also criticized the methodology of studies examining the bilingual advantage such as a failure to match participants on demographic variables. Furthermore, they pointed out the publication bias in the field that makes it considerably more likely for positive findings to be published over negative or null findings (De Bruin et al., 2015). This suggests that the positive findings that eventually get published may be atypical and not representative of how bilingualism affects executive functioning. As a result, they concluded that the phenomenon did not extent beyond random, atypical findings.

Samuel and colleagues (2018) also failed to find evidence for the bilingual advantage in any age group upon review of the literature. In fact, their results appeared to support a bilingual *disadvantage* in executive functioning. Additionally, when a group of East Asian young adults were compared to their Western counterparts, the former group outperformed the latter on measures of executive functioning regardless of bilingual status. Given such disconfirming evidence as well as previous literature suggesting that an East Asian background may contribute to group differences usually attributed to the bilingual advantage (Abutalebi et al., 2015; Ong et al., 2017), they concluded that culture could be a confounding factor in examining the bilingual advantage.

In other words, the group differences often attributed to bilingualism may be due to cultural effects. There is also evidence that suggests a substantial publication bias in this field of research favoring positive results over negative or mixed findings (De Bruin et al., 2015). A recent meta-analysis showed that the small bilingual advantage disappeared after publication bias was corrected for (Lehtonen et al., 2018). Therefore, the often invisible influence of publication bias casts further doubt on the reliability and magnitude of bilingual advantage.

Evidence of a Bilingual Disadvantage

The strongest evidence against the bilingual advantage comes from investigators who found that bilinguals perform *more poorly* on certain cognitive tasks. Bilinguals tend to exhibit lower performance than their monolingual counterparts on measures that require lexical access (Michael & Gollan, 2005). A study by Roberts and colleagues (2002) found that bilinguals were able to name fewer pictures on the Boston Naming Test (BNT) than monolinguals despite being highly proficient in English. These findings were replicated by Gollan and colleagues (2005) who also found that bilinguals made more errors in picture naming than monolinguals. Bilinguals also named pictures more slowly, which may be due to experiencing tip-of-the-tongue (TOT) retrieval failures more frequently than their monolingual peers (Gollan & Silverberg, 2001).

Further, there is evidence that bilinguals have lower verbal fluency than monolinguals. Gollan and colleagues (2002) found that Spanish-English bilinguals performed more poorly than English monolinguals on semantic fluency, category fluency and proper name fluency tasks. Further, the group difference was most pronounced on the sematic fluency task. This bilingual disadvantage in verbal fluency remains evident even when bilinguals are tested in their dominant or first language (Ivanova & Costa, 2008). Roselli and colleagues (2000), on the other hand, found that bilinguals and monolinguals performed equally on all verbal fluency tasks with the exception of the semantic fluency task where the bilinguals performed more poorly.

Clinical Relevance of the Debate on Bilingual (Dis)Advantage

While the controversy around the net effect of bilingualism on cognitive functioning has inspired important studies within cognitive neuroscience and psycholinguistics, its relevance to applied settings remains unclear. In clinical assessment, the main practical issue is the examinee's level of *proficiency in the language in which tests are administered*. The vast majority of cognitive tests used in North America were normed on (sometimes monolingual) native speakers of English (NSE). In fact, NSE status is an implicit assumption during neuropsychological assessment. Violating this assumption creates significant barriers to the meaningful evaluation of neurocognitive functioning (Cardenas, Villavicencio & Pavuluri, 2017; Mindt et al., 2008). The subtle influence of being proficient in other languages *in addition to* English pales in comparison.

Native vs Non-Native Speakers of English

Since tests were developed for and normed on NSEs, the extent to which they can be applied to examinees with limited English proficiency (LEP) is largely unknown. The term LEP refers to individuals who learned English later in life, outside the sensitive window of language acquisition (age 12–15; Johnson & Newport, 1989; Lenneberg, 1967; Sakai, 2005). LEP is a descriptive label that captures a significant deviation in language proficiency from NSE. It is more precise than commonly used alternative terminology such as “English as a second language” or “English language learners” (Ali et al., 2020).

In contrast to the abundance of research on bilingualism, the literature on cognitive profiles associated with LEP appears to be limited. Grundy and Timmer (2016) noted in their meta-analysis on working memory and bilingualism that the studies they included sometimes tested the participants in one of the languages they spoke or both. This variability in operationalizing and studying bilingualism indicates that the NSE-LEP distinction is not a salient variable in bilingual research, even though it is an important one in a clinical context.

Even in clinical research, most of the literature is focused on examining the neuropsychological performance of bilinguals, without making LEP an explicit target of the investigation. Nevertheless, the sporadic evidence is slowly consolidating into a coherent pattern of deficits associated with LEP on performance based psychometric testing. The effect of LEP on test performance appears to be mediated by a number of factors, some of which are related to the instruments (level of verbal mediation, cognitive domain, type of test; Gooding, Cole & Hamberger, 2018), and some of which are related to demographic factors (age, level of education, country of origin; Boone et al., 2007).

Level of Verbal Mediation

Based on whether NSE status is critical to performing a given task, psychometric tests can be divided into two categories: low verbal mediation (LVM) and high verbal mediation (HVM). In the present work, LVM and HVM tests were defined based on previous research on neurocognitive profiles of LEP reviewed in the next section. Tests with LVM still require a basic proficiency in the language of administration, but once examinees comprehend the instructions, their verbal skills are assumed to be orthogonal to their test performance. Examples of LVM tests include visual memory and motor speed.

In contrast, on tests with HVM native-level language proficiency is thought to be a necessary condition to demonstrating one's true ability level on the target construct (Erdodi, Jongsma, & Issa, 2017). Measures developed to diagnose aphasia are automatically considered HVM measures, given that aphasia refers to a focal deficit in language skills that is conceptually similar to LEP. Examples of HVM tests include verbal memory, measures of receptive and expressive language or verbal reasoning.

Naturally, LEP is expected to have a significant deleterious effect on tests with HVM, but less so on tests with LVM. The empirical evidence is largely consistent with this prediction (Boone et al., 2007; Erdodi, Jongsma & Issa, 2017; Gooding et al., 2018; Kisser et al, 2012). Given the lack of scientific consensus on operationalizing verbal mediation, for the purpose of this study tests were classified based on previous empirical evidence or, in the absence of that, on rational grounds.

The Neurocognitive Profiles of LEP

Individuals with LEP perform worse than NSEs on neuropsychological tests with HVM. Boone and colleagues (2007) reported that NSEs performed better on the BNT than people with

LEP. Subsequent investigations found that LEP was associated with deficits in object naming (i.e., performance on the BNT and Auditory & Visual Naming Test; Gooding et al., 2018) as well as category and letter fluency (Boone et al, 2007; Erdodi, Jongsma & Issa, 2017; Kisser et al, 2012). These findings suggest that in comparison to NSEs, individuals with LEP perform poorly on neuropsychological tests with HVM.

In contrast, studies comparing LEP and NSE on measures of attention and processing speed produced either null findings or higher performance in NSE, depending on the test used (Walker et al., 2010). Harris, Wagner and Munro Cullom (2007) demonstrated equivalent performance between NSE and LEP on the Digit Symbol Coding (CD) subtest of the Wechsler Adult Intelligence Scale (WAIS). These results were replicated by later studies (Erdodi, Jongsma & Issa, 2017; Walker et al., 2010). Interestingly, Harris and colleagues found a negative relationship between level of education and test performance within the LEP sample. Equivalent performance was also reported on the Rey 15-Item Test (Rey-15; Salazar et al., 2007), a free-standing performance validity test (PVT) based on visual memory.

The NSE advantage re-emerged on other measures of attention and processing speed. LEP status was associated with lower performance on the Digit Span subtest of the WAIS, which measures auditory attention and working memory (Boone et al., 2007; Lozano-Ruiz et al.; Walker et al., 2010). Interestingly, individuals with LEP also demonstrated poor performance relative to NSE individuals on Symbol Search, an analog measure of processing speed on the WAIS. LEP was associated with lower performance on Trails A and the Stroop B (color naming trial), even though these are commonly considered tests with LVM (Razani et al., 2007; Kisser et al., 2012).

With regards to visuospatial abilities, the findings appear more mixed. NSE and LEP groups appear to have equivalent performance on the Clock Drawing Test (CDT), a measure with LVM (Borson et al., 1999; Erdodi, Nussbaum, et al., 2017; Menon et al., 2012). A surprising finding was that on the copy trial of the Rey-Osterrieth Complex Figure Test (RCFT), a LVM task, the individuals with LEP demonstrated *higher* performance than NSEs (Boone et al., 2007). This is the only neuropsychological test on which an LEP advantage emerged, although the reasons for this are not clear.

On tests of executive functioning, NSEs perform better than the LEP group. Razani et al. (2007) found that NSEs outperformed the LEP group on the Stroop C (color-reading interference condition) and Trails B.

Summary of Previous Research

A cursory review of the literature on cognitive profiles associated with bilingualism and LEP revealed divergent findings. Results varied as a function of how the opposite of "monolingual NSE" was defined (i.e., "bilingual" or "LEP"), level of verbal mediation, cognitive domains, tests, and even subtests. The simple fact that the arc of the review spans from discussing evidence of a bilingual *advantage* in executive functioning to a *deficit* within the same cognitive domain among individuals with LEP emphasizes the importance of criterion grouping. Although the research on cognitive profiles associated with LEP is far from conclusive, it has produced more consistent findings than the research on bilingualism. Therefore, the present study focuses on the *NSE vs LEP* contrast to examine the pattern of strengths and weaknesses associated with LEP from an applied clinical perspective.

Present Study

This study sought to continue the existing line of systematic investigations (Kisser et al., 2012; Razani et al., 2007) of neurocognitive profiles associated with LEP. The performance of cognitively healthy participants (half NSEs, and half individuals with LEP) was compared on a battery of neuropsychological tests, evenly split between LVM and HVM instruments. A significant NSE advantage was predicted on tests with HVM (medium-large effects). No difference between the NSE and LEP sample was predicted on tests with LVM. Table 1 lists the tests administered.

CHAPTER 2

Methods

Participants

This study was based on a previous research project designed to examine the differential failure rate on performance validity tests (PVTs) as a function of LEP status and level of verbal mediation (Abeare, An, et al., in press; Abeare, Cutler, et al., An et al., 2019; Erdodi et al., 2020). In contrast, the present investigation was focused on between-group differences on measures of cognitive ability rather than PVTs. Participants were undergraduate and graduate students recruited from the University of Windsor Participant Pool, Centre for English Language Development (CELD), and Windsor International Student Email List (WISEL). The students were compensated with either extra credit or $20 if they were ineligible for extra credit.

Participants were excluded if they had a current diagnosis of a major psychiatric or neurological disorder, developmental disability, or serious medical illness that would significantly affect performance on testing. Individuals were assigned to either the LEP or NSE group based on their scores on the English Language Proficiency rating scales (speaking, understanding, reading) of Language Experience and Proficiency Questionnaire (LEAP-Q). The study was approved by the University of Windsor's Research Ethics Board. APA ethical guidelines regulating research with human participants were followed throughout the process. A total of 80 participants were included in the study (40 LEP, 40 NSE). Both the LEP and NSE samples included bilingual participants. The LEP sample included individuals of varying cultural and linguistic backgrounds.

Measures

English language proficiency was assessed using the Language Experience and Proficiency Questionnaire (LEAP-Q; Marian et al., 2007). A brief battery of neuropsychological measures examining several cognitive domains (e.g., memory, executive function, attention, processing speed) was administered in order to mimic a real-world assessment situation. Additionally, the battery was strategically selected to achieve a balance between tests with HVM and LVM (see Table 1 for a comprehensive list of the measures used). A brief demographic questionnaire was used to collect relevant demographic information (e.g., age, gender, SES).

Procedure

Testing was conducted in a quiet environment by trained research assistants (RAs). The testing procedure, compensation, risks and benefits were explained to the participants by the primary investigator (PI). After informed consent was obtained, the PI left the room, and the battery of neuropsychological tests was administered by the RAs in counterbalanced order. Both the RAs and participants were blinded to the hypotheses of the study in order to avoid biases in test administration.

Data Analysis

Descriptive statistics such as means and standard deviations were calculated and reported for demographic variables such as age, education, and age of English language acquisition. Independent *t*-tests (two-tailed) were used to compare the LEP and NSE groups on each of the neuropsychological measures. Group averages were also compared to normative data. The effect size was expressed in Cohen's *d*. Bonferroni corrections were not used to correct for multiple comparisons due to several methodological counter-indications such as suppressing the discovery rate in new lines of investigation and artificially reduced statistical power (Moran, 2003; Perneger, 1998). Moreover, Bonferroni corrections address the question of whether the universal null hypothesis is true, and not the specific research questions at hand (Savitz & Olshan, 1995). As a result, computing effect sizes provide a better indication of the relevant experimental effects in this study, and their practical significance (Nakagawa, 2004). Further, all the comparisons were *a priori* planned, and the cost of false discovery is low, given that this was an exploratory study.

CHAPTER 3

Results

Demographic Differences

There were no between-group differences on age, lateral dominance and self-reported depression or anxiety (Table 2). However, a larger proportion of the NSE sample was female (85.0% versus 57.5%). Also, the LEP sample had a higher level of education (medium-large effect). Naturally, the NSE sample learned English at an earlier age (extremely large effect).

Between-Group Differences in Neurocognitive Profiles

Tests with HVM

The NSE outperformed the LEP sample on most tests with HVM, with three exceptions: Digit Span, Rey WRT and the D-KEFS Word Reading trials (*p*: .099-.528). Among the significant contrasts, the largest effects were observed on the BNT-15 (*d*: 3.14-3.17; extremely large), followed by the CIM (*d* = 1.92, very large), category fluency (*d* = 1.62, very large), and single-word reading (*d* = 1.58, very large). The deficit associated with LEP on the D-KEFS Stroop subtest was commensurate with that observed on the Color Naming subtest (*d*: 1.09-1.12, large effects). On the WCT as a continuous variable, LEP was associated with a lower accuracy (*d* = 0.48, medium effect) and longer time-to-completion (*d* = 0.77, large effect) score (Table 3). Most significant contrasts for mean differences were also associated with unequal variance. In the majority of cases, the LEP sample produced *higher* within-group variability. However, on two of the tests (letter and emotion word fluency) participants with LEP produced *lower* SDs.

Tests with LVM

No significant difference emerged on most of the LVM tests (Table 4). Notable exceptions include the Clock Drawing Task (large effect), RCFT Recognition trial (medium effect), TMT-B (medium effect) and TOMM-1 (medium effect). Worth noting, the LEP group scored *higher* on the TOMM-1. Participants with LEP also produced more variable scores on Clock Drawing and the Rey-15.

Between-Group Differences from the Normative Mean

Tests with HVM

NSEs produced lower means than the normative sample on letter fluency (*d* = 0.70, large effect) and CIM (*d* = 0.42, medium effect). On the D-KEFS Stroop, however, the NSE group mean was significantly *higher* than the normative mean (*d* = 0.93, large effect). There was no significant difference on the other HVM tests (Table 5).

In contrast, the LEP sample mean was lower than the normative mean on category and letter fluency, CIM, D-KEFS Color Naming and Stroop condition, and single word reading. Effect sizes ranged from low-medium to extremely large (*d:* 0.37-2.70). No significant differences emerged on the WAIS-III Digit Span and D-KEFS Word Reading.

Tests with LVM

NSE performed below the normative sample on the TMT-A (medium effect). However, they performed above the normative mean on Coding and Symbol Search (medium/large effect). None of the other contrasts were significant (Table 6). The advantage relative to the normative sample on Coding and Symbol Search was maintained in the LEP sample (medium effect). However, participants with LEP scored below the normative mean on all RCFT trails (medium effect), and TMT-B (large effect).

CHAPTER 4

Discussion

Main Findings

The objective of the present study was to contribute to the existing body of research on the neurocognitive profiles associated with LEP. It was hypothesized that the LEP and NSE groups would be equivalent on LVM tests, but an LEP disadvantage would emerge on the HVM measures. The data were broadly consistent with these hypotheses, with a few surprising exceptions, which suggests that the HVM/LVM distinction may be more complex than it first appears. The LEP disadvantage on HVM tests was most pronounced on measures used to diagnose aphasia (e.g. Animals, BNT-15, CIM), a free-standing PVT based on word recognition memory (WCT), and measures of processing speed and executive functioning (i.e. D-KEFS Color, D-KEFS Stroop). The LEP and NSE groups were equivalent on LVM measures of processing speed (i.e. WAIS-III Coding and Symbol Search, TMT-A), measures of visuospatial skills (i.e. RCFT Copy, RCFT Immediate Recall, RCFT Delayed Recall) and a free-standing PVT based on visual memory (Rey-15).

Unexpected Findings

The NSE sample outperformed the LEP sample on most HVM tests with the exception of the Digit Span, Rey WRT and the D-KEFS Word Reading. These results contradict previous findings showing an LEP disadvantage on the WAIS Digit Span (Boone et al., 2007; Walker et al., 2010), and may be explained by the differences in sample characteristics. The aforementioned studies were based on clinical samples. Consequently, neuropsychological deficits and LEP may have had a cumulative deleterious effect. In contrast, this study was based on cognitively intact university students. Therefore, the effect of LEP on test performance was observed without the confounding effect of cognitive or emotional problems. Further, both samples had higher levels of education than previous studies, and the LEP group had significantly higher levels of education than the NSE group. Thus, these unusual findings may be due to these sample differences, given level of education is positively correlated with performance on cognitive testing (Ganguli et al., 2010; Walker et al., Shores, 2009).

The pattern of between-group performance on the D-KEFS (LEP < NSE on Color Naming and Stroop, LEP = NSE on Word Reading) suggests that not all subtests are equally susceptible to the deleterious effect of LEP. This finding could be explained by the fact that bilinguals activate the phonology of both languages during naming tasks, particularly when they are being tested in a language in which they have lower proficiency (Macizo, 2016). This phenomenon is posited to raise the difficulty level of a task for examinees with LEP, as they are confronted with an *inhibition task* even during the baseline trials of the test. The effect is expected to be stronger on the color naming task compared to the word reading task, as the target stimulus (patch of color) is more likely to automatically elicit the response (name of the color) in the dominant language, which the examinee with LEP has to inhibit, and choose the English name for it instead.

In other words, for examinees with LEP, the *interference* task on the D-KEFS starts with Color Naming. This additional processing requirement due to the need to manage to the co-activation of two languages may explain the slower completion time during the D-KEFS Color Naming and Stroop tasks. In combination with the equivalent performance on the Rey WRT (a verbal recognition memory task), these findings suggest that the rationally based HVM/LVM classification used in this study had inconsistent empirical support (Kisser et al., 2012). Instead of judging the level of verbal mediation involved in a given test on purely rational grounds, the construct should be evaluated using objective, performance-based data.

On the LVM tests, there was no difference in performance between groups, with four exceptions: the Clock Drawing Task, RCFT Recognition trial, TMT-B and the TOMM-1. Similarly, these findings run counter to previous research that demonstrated equivalent performance between LEP and NSE groups on visuospatial measures such as the Clock Drawing Task (Erdodi, Jongsma & Issa, 2017). This could potentially be accounted for by cultural factors, given that in the previous study, both the LEP and NSE groups consisted of English-Arabic bilinguals. The current study, on the other hand, included participants of varying cultural backgrounds. Additionally, although at face value, the CDT is a LVM visuospatial measure, it also requires the *comprehension* of the test instructions provided in English and maintaining them in the short-term *verbal memory*, which is likely a HVM task component. The lexical format itself may vary across languages (i.e., “eleven and ten” versus “ten after eleven”), or the salience of the analog clock itself may vary across cultures. It is, however, possible that these unexpected LVM findings are spurious results due to the lack of correction for multiple comparisons.

Razani et al. (2007) also found an LEP disadvantage on TMT-B, which is consistent with the results of the current study. Stålhammar and colleagues (2020) hypothesized that the mastery of the alphabet of a non-native language would not guarantee native recital speed even on nonvocalized tasks such as the TMT-B. They also hypothesized that differences in alphabetical and/or numerical symbols across languages could affect LEP performance on the TMT-B. Differences in academic traditions across countries (e.g., relative emphasis on written symbols versus oral communication) may have been an additional confounding variable.

Surprisingly, the LEP sample outperformed the NSE sample on the TOMM-1. Previous research shows that in linguistically and culturally diverse samples, those with high levels of education fail the TOMM at lower rates compared to less educated individuals (Nijdam-Jones et al., 2019; Rivera et al. 2015). In the current sample, the LEP group has significantly more education than the NSE group, which could have amplified this effect.

NSE versus LEP Compared to Normative Data

Among HVM tests, the NSE subsample performed more poorly on letter fluency and CIM when compared to the normative sample. This unexpected finding may be due to the inclusion of bilingual participants in the NSE sample. Previous research shows that bilinguals have a disadvantage on tests with HVM compared to monolinguals, even when they are highly proficient in English (Michael & Gollan, 2005; Roberts et al., 2002). Thus, the presence of bilingual participants in the NSE sample may have had a significant impact on the results given that the CIM normative sample is composed of monolingual English speakers (Borod et al., 1980). NSEs performed above the normative sample on the D-KEFS Stroop, but showed no difference in performance in the D-KEFS Color Naming and Word Reading conditions. This also may be due to the presence of bilinguals in the sample, given that there is evidence for a bilingual advantage in executive functioning (Hilchey & Klein, 2011). This would have a greater effect on the D-KEFS Stroop because it is primarily a measure of executive functioning, whereas the D-KEFS Color Naming and Word Reading conditions are measures of attention and processing speed.

Consistent with expectations, the LEP subsample performed very poorly compared to the normative sample on most of the HVM tests, specifically on single word reading, category and letter fluency, CIM, D-KEFS Color Naming and Stroop condition. Additionally, past research shows an LEP disadvantage on category and letter fluency, as well as color naming and the Stroop task (Kisser et al., 2012; Razani et al., 2007).

On LVM tests, NSEs performed below the normative sample on TMT-A, although they performed above the normative mean on Coding and Symbol Search. High performance on Coding and Symbol Search could have been due to the high level of education (Heaton et al., 2004). This is consistent with evidence showing that high levels of education is associated with higher performance on the WAIS Processing Speed Index, which is composed of WAIS Coding and Symbol Search (Heaton, Taylor, & Manly, 2003). Similarly, students volunteering for academic research who demonstrated high test taking effort routinely perform above the normative mean (Abeare et al., 2019; Hurtubise et al., 2020). The lower performance on TMT-A may be due to differences in sample characteristics, as past research suggests that the TMT is susceptible to the influence of culture (Fernandez & Marcopulos, 2008).

Participants with LEP performed above the normative mean on Coding and Symbol Search. However, they performed poorly on RCFT, TMT-A and TMT-B. Lower LEP performance on TMT-A and TMT-B compared to NSE is consistent with findings from past studies and may be attributable to a lack of automaticity with the 26-letter Latin alphabet as hypothesized in past literature (Kisser et al., 2012; Razani et al., 2007; Stålhammar et al., 2020).

Clinical Implications of the Findings

The findings showed that the majority of the HVM tests likely underestimated true cognitive ability in individuals with LEP. This suggests that low scores on such measures should not be interpreted as evidence of acquired cognitive impairment in individuals with LEP, since poor performance may simply be a natural consequence of LEP status. Further, an LEP disadvantage was observed even on some LVM tests, suggesting that the deleterious effects of LEP is not limited to tests with HVM.

Notably, LEP suppressed performance on a single-word reading test, which is a measure commonly used by neuropsychologists to estimate premorbid functioning (Bright et al., 2002; Bright & van der Linde, 2020; Schretlen et al., 2005). Thus, this test should not be used for that purpose in LEP patients, as it may systematically underestimate their baseline cognitive ability. Finally, there was no evidence for an LEP-specific disadvantage on the Rey-15, a free-standing PVT, which suggests that this instrument may still be used in individuals with LEP. This finding runs contrary to some previous reports (Ali et al., 2020; Salazar et al., 2007; Strutt et al., 2011). It is, however, notable that some of these studies used samples with lower levels of education, which could have confounded the findings (Salazar et al., 2007; Strutt et al., 2011).

Strengths

This study's main strength is comparing neuropsychological test performance in individuals with LEP by to NSEs, as well as comparing both groups to the normative sample.

The study also provides a representative sample of commonly used neuropsychological tests. Finally, study participants were healthy controls at their cognitive peak with no incentive to appear impaired. Therefore, it provides a snapshot of the typical differences in cognitive profiles between NSE and LEP.

Limitations

There are, however, some limitations to be considered. Due to demographic and geographic limitations, all participants were students at a single university from a mid-sized Canadian city. Thus, results may not generalize to other populations given that past research shows geographic and demographic differences in performance on cognitive tests (Lichtenstein et al., 2019; McDaniel, 2006).

Additionally, the sample was comprised of cognitively healthy, high functioning, highly educated individuals. Genuine neuropsychiatric disorders may interact with LEP, compounding its deleterious effect, thus the findings must be replicated in clinical patients with LEP instead of extrapolating the findings from controls. Another limitation of the study was the significant difference in education and gender between groups. Specifically, the LEP group had more years of education, and NSE group had a higher proportion of female participants, both of which could have confounded the results of the study. Many of the participants in the NSE group were bilinguals, and as such the criterion groups were contaminated, making it more difficult to isolate the effect of LEP.

Finally, the LVM/HVM distinction lacks a solid epistemological basis. The relatively high number of counterintuitive findings even within the present investigation serves as a reminder of the unreliability of theoretically derived classification of instruments into these categories. Until the evidence base on level of verbal mediation in general and for individual instruments consolidates, this construct should be used with caution in LEP research. In the interim, investigators are cautioned against using HVM/LVM as the main grouping variable. Future research comparing a monolingual native speaker group, a bilingual NSE group and an LEP group would help expand our current understanding of this test parameter.

Conclusion

The findings of the current study contribute to the existing body of research on neurocognitive profiles of LEP by comparing LEP and NSE performance on measures with differing levels of verbal mediation. The results could potentially aid neuropsychologists in deciding which measures to use in assessing patients with LEP and provide evidence based interpretative guidelines for estimating cognitive deficits. Future research is needed to replicate results in clinical populations, different bilingual or LEP samples, as well as examine different tests representing the HVM/LVM construct.

REFERENCES

Abeare, C. A., An, K., Tyson, B., Holcomb, M., Cutler, L., May, N., & Erdodi, L. A. (In Press). The emotion word fluency test as an embedded validity indicator, alone and in multivariable verbal fluency validity composite. *Applied Neuropsychology: Child, 0*(0), 00-00. Advance online publication.

Abeare, K., Cutler, L., An, K. Y., Razvi, P., Holcomb, M., & Erdodi, L. A. (In Press). The Boston naming test – short form: revised performance validity cutoffs and proposed clinical classification ranges. *Cognitive and Behavioral Neurology, 0*(0), 00-00. Advance online publication.

Abeare, C. A., Freund, S., Kaploun, K., McAuley, T., & Dumitrescu, C. (2017). The emotion word fluency test (EWFT): Initial psychometric, validation, and physiological evidence in young adults. *Journal of Clinical and Experimental Neuropsychology*, *39*(8), 738-752. https://doi.org/10.1080/13803395.2016.1259396

Abeare, C., Messa, I., Whitfield, C., Zuccato, B., Casey, J., Rykulski, N., & Erdodi, L. (2019). Performance validity in collegiate football athletes at baseline neurocognitive testing. *The Journal of Head Trauma Rehabilitation*, *34*(4), E20-E31. https://doi.org/10.1097/htr.0000000000000451

Abutalebi, J., Guidi, L., Borsa, V., Canini, M., Della Rosa, P. A., Parris, B. A., & Weekes, B. S. (2015). Bilingualism provides a neural reserve for aging populations. *Neuropsychologia*, *69*, 201-210. https://doi.org/10.1016/j.neuropsychologia.2015.01.040

Adesope, O. O., Lavin, T., Thompson, T., & Ungerleider, C. (2010). A systematic review and meta-analysis of the cognitive correlates of bilingualism. *Review of Educational Research, 80,* 207–245. https://doi.org/10.3102%2F0034654310368803

Alansari, B., & Baroun, K. (2004). Gender and cultural performance differences on the stroop color and word test: A comparative study. *Social Behavior and Personality: An International Journal*, *32*(3), 235-245. https://doi.org/10.2224/sbp.2004.32.3.235

Alladi, S., Bak, T. H., Duggirala, V., Surampudi, B., Shailaja, M., Shukla, A. K., Chaudhuri, J. R., & Kaul, S. (2013). Bilingualism delays age at onset of dementia, independent of education and immigration status. *Neurology*, *81*(22), 1938-1944. https://doi.org/10.1212/01.wnl.0000436620.33155.a4

An, K. Y., Charles, J., Ali, S., Enache, A., Dhuga, J., & Erdodi, L. (2019). Reexamining performance validity cutoffs within the complex ideational material and the Boston naming test-short form using an experimental malingering paradigm. *Journal of Clinical and Experimental Neuropsychology, 41*(1), 15-25. https://doi.org/10.1080/13803395.2018.1483488

Ansaldo, A. I., Marcotte, K., Scherer, L., & Raboyeau, G. (2008). Language therapy and bilingual aphasia: Clinical implications of psycholinguistic and neuroimaging research. *Journal of Neurolinguistics*, *21*(6), 539-557. https://doi.org/10.1016/j.jneuroling.2008.02.001

Antoniou, M. (2019). The advantages of bilingualism debate. *Annual Review of Linguistics*, *5*, 395-415. https://doi.org/10.1146/annurev-linguistics-011718-011820

Bialystok, E. (2011). Reshaping the mind: the benefits of bilingualism. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *65*(4), 229. https://doi.apa.org/doi/10.1037/a0025406

Bialystok, E., Craik, F. I., & Freedman, M. (2007). Bilingualism as a protection against the onset of symptoms of dementia. *Neuropsychologia*, *45*(2), 459-464. https://doi.org/10.1016/j.neuropsychologia.2006.10.009

Boone, K. B., Victor, T. L., Wen, J., Razani, J., & Ponton, M. (2007). The association between neuropsychological scores and ethnicity, language, and acculturation variables in a large patient population. *Archives of Clinical Neuropsychology*, *22*, 355-365. https://doi.org/10.1016/j.acn.2007.01.010

Borod, J. C., Goodglass, H., & Kaplan, E. (1980). Normative data on the Boston diagnostic aphasia examination, parietal lobe battery, and the Boston naming test. *Journal of Clinical Neuropsychology, 2*, 209–215. https://doi.org/10.1080/01688638008403793

Borson, S. O. O., Brush, M., Gil, E., Scanlan, J., Vitaliano, P., Chen, J., Cashman, J., Sta Maria, M. M., Barnhart, R., & Roques, J. (1999). The clock drawing test: Utility for dementia detection in multiethnic elders. *Journals of Gerontology Series A: Biomedical Sciences and Medical Sciences*, *54*(11), M534-M540. https://doi.org/10.1093/gerona/54.11.M534

Bright, P., Jaldow, E. L. I., & Kopelman, M. D. (2002). The national adult reading test as a measure of premorbid intelligence: A comparison with estimates derived from demographic variables. *Journal of the International Neuropsychological Society: JINS*, *8*(6), 847. https://doi.org/10.1017/s1355617702860131

Bright, P., & van der Linde, I. (2020). Comparison of methods for estimating premorbid intelligence. *Neuropsychological Rehabilitation*, *30*(1), 1-14. https://doi.org/10.1080/09602011.2018.1445650

Cardenas, A., Villavicencio, L., & Pavuluri, M. (2017). Considerations in the neuropsychological evaluation and treatment of children with limited English proficiency. *The Permanente Journal*, *21*. https://doi.org/10.7812/TPP/16-149

Coffey, D. M., Marmol, L., Schock, L., & Adams, W. (2005). The influence of acculturation on the Wisconsin card sorting test by Mexican Americans. *Archives of Clinical Neuropsychology*, *20*(6), 795-803. https://doi.org/10.1016/j.acn.2005.04.009

Crystal, D. (1997). English as a global language. Cambridge, England: Cambridge University Press. https://doi.org/10.1017/CBO9780511486999

de Bruin, A., Treccani, B., & Della Sala, S. (2015). Cognitive advantage in bilingualism: An example of publication bias?. *Psychological Science*, *26,* 99–107. https://doi.org/10.1177%2F0956797614557866

Delis, D. C., Kaplan, E., & Kramer, J. H. (2001). *D-KEFS: Examiner's manual.* The Psychological Corporation. https://doi.org/10.1037/t15082-000

Erdodi, L. A., Jongsma, K. A., & Issa, M. (2017). The 15-item version of the Boston naming test as an index of English proficiency. *The Clinical Neuropsychologist*, *31*(1), 168-178. https://doi.org/10.1080/13854046.2016.1224392

Erdodi, L. A., Nussbaum, S., Sagar, S., Abeare, C. A., & Schwartz, E. S. (2017). Limited English proficiency increases failure rates on performance validity tests with high verbal mediation. *Psychological Injury and Law*, *10*(1), 96-103. https://doi.org/10.1007/s12207-017-9282-x

Erdodi, L. A., Sabelli, A. G., An, K. Y., Hastings, M., McCoy, C., & Abeare, C. A. (2020). Introducing a five-variable psychiatric screener based on the visual analog scale (V-5). *Psychology & Neuroscience.* https://doi.org/10.1037/pne0000201

Eurobarometer, S. (2012). Europeans and their languages. *European Commission*.

Eurostat. (2016). Foreign language skills statistics [Data file]. Retrieved from https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Foreign_language_skills_statistics

Fernandez, A. L., & Marcopulos, B. A. (2008). A comparison of normative data for the trail making test from several countries: Equivalence of norms and considerations for interpretation. *Scandinavian Journal of Psychology*, *49*(3), 239-246. https://doi.org/10.1111/j.1467-9450.2008.00637.x

Freedman, M., Alladi, S., Chertkow, H., Bialystok, E., Craik, F. I., Phillips, N. A., ... & Bak, T. H. (2014). Delaying onset of dementia: Are two languages enough?. *Behavioural Neurology*, *2014*. https://doi.org/10.1155/2014/808137

Ganguli, M., Snitz, B. E., Lee, C. W., Vanderbilt, J., Saxton, J. A., & Chang, C. C. H. (2010). Age and education effects and norms on a cognitive test battery from a population-based cohort: The Monongahela–Youghiogheny healthy aging team. *Aging and Mental Health*, *14*(1), 100-107. https://doi.org/10.1080/13607860903071014

Gollan, T. H., Montoya, R. I., & Werner, G. A. (2002). Semantic and letter fluency in Spanish-English bilinguals. *Neuropsychology*, *16*(4), 562. https://doi.org/10.1037/0894-4105.16.4.562

Gollan, T. H., & Silverberg, N. B. (2001). Tip-of-the-tongue states in Hebrew–English bilinguals. *Bilingualism: Language and Cognition*, *4*(1), 63-83. https://doi.org/10.1017/S136672890100013X

Goodglass, H., & Kaplan, E. (1972). *The assessment of aphasia and related disorders*. Lea & Febiger.

Gooding, A., Cole, J. R., & Hamberger, M. J. (2018). Assessment of naming in non-native English speakers with epilepsy. *Journal of the International Neuropsychological Society*, *24*(10), 1057-1063. https://doi.org/10.1017/s1355617718000632

Grosjean, F. (1994). Individual bilingualism. *The Encyclopedia of Language and Linguistics*, *3*, 1656-1660.

Grundy, J. G., & Timmer, K. (2016). Bilingualism and working memory capacity: A comprehensive meta-analysis. *Second Language Research, 33*, 325–340. https://doi.org/10.1177/0267658316678286

Harris, J. G., Wagner, B., & Munro Cullum, C. (2007). Symbol vs. digit substitution task performance in diverse cultural and linguistic groups. *The Clinical Neuropsychologist*, *21*(5), 800-810. https://doi.org/10.1080/13854040600801019

Heaton, R. K., Taylor, M. J., & Manly, J. (2003). Demographic effects and use of demographically corrected norms with the WAIS-III and WMS-III. In D. S. Tulsky, D. H. Saklofske, R. K. Heaton, R. Bornstein, M. F. Ledbetter, G. J. Chelune, R. J. Ivnik, & A. Prifitera (Eds.), *Clinical interpretation of the WAIS-III and WMS-III* (pp. 181-210). Academic Press. https://doi.org/10.1016/B978-012703570-3/50010-9

Heaton, R., Miller, W., Taylor, M., & Grant, I. (2004). *Revised comprehensive norms for an expanded halstead-reitan battery: Demographically adjusted neuropsychological norms for African American and Caucasian adults*. Psychological Assessment Resources, Inc.

Hilchey, M. D., & Klein, R. M. (2011). Are there bilingual advantages on nonlinguistic interference tasks? Implications for the plasticity of executive control processes. *Psychonomic Bulletin & Review, 18*, 625–658. https://doi.org/10.3758/s13423-011-0116-7

Hurtubise, J., Baher, T., Messa, I., Cutler, L., Shahein, A., Hastings, M., ... & Erdodi, L. A. (2020). Verbal fluency and digit span variables as performance validity indicators in experimentally induced malingering and real world patients with TBI. *Applied Neuropsychology: Child, 9*(4), 337-354. https://doi.org/10.1080/21622965.2020.1719409

Ivanova, I., & Costa, A. (2008). Does bilingualism hamper lexical access in speech production?. *Acta Psychologica, 127*(2), 277-288. http://dx.doi.org/10.1016/j.actpsy.2007.06.003

Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, *21*(1), 60–99. https://doi.org/10.1016/0010-0285(89)90003-0

Kisser, J. E., Wendell, C. R., Spencer, R. J., & Waldstein, S. R. (2012). Neuropsychological performance of native versus non-native English speakers. *Archives of Clinical Neuropsychology*, *27*(7), 749-755. https://doi.org/10.1093/arclin/acs082

Kowoll, M. E., Degen, C., Gladis, S., & Schröder, J. (2015). Neuropsychological profiles and verbal abilities in lifelong bilinguals with mild cognitive impairment and Alzheimer's disease. *Journal of Alzheimer's Disease*, *45*(4), 1257-1268. http://dx.doi.org/10.3233/JAD-142880

Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine*, *16*(9), 606-613. https://doi.org/10.1046/j.1525-1497.2001.016009606.x

Lehtonen, M., Soveri, A., Laine, A., Järvenpää, J., De Bruin, A., & Antfolk, J. (2018). Is bilingualism associated with enhanced executive functioning in adults? A meta-analytic review. *Psychological Bulletin*, *144*(4), 394. https://doi.org/10.1037/bul0000142

Lenneberg, E. H. (1967). The biological foundations of language. *Hospital Practice*, *2*(12), 59-67. https://doi.org/10.1080/21548331.1967.11707799

Lezak, M. D., Howieson, D. B., Loring, D. W., & Fischer, J. S. (2004). *Neuropsychological assessment.* Oxford University Press.

Lichtenstein, J. D., Greenacre, M. K., Cutler, L., Abeare, K., Baker, S. D., Kent, K. J., Ali, S., & Erdodi, L. A. (2019). Geographic variation and instrumentation artifacts: In search of confounds in performance validity assessment in adults with mild TBI. *Psychological Injury and Law*, *12*(2), 127-145. https://doi.org/10.1007/s12207-019-09354-w

Lozano-Ruiz, A., Leonard, B., Daugherty, J. C., Fernandez-Fillol, C., Puente, A. E., & Perez-Garcia, M. Language and culture differences in working memory: A cross-cultural study through the embraced digit span task. http://dx.doi.org/10.13140/RG.2.2.32918.09284

Mack, W. J., Freed, D. M., Williams, B.W., & Henderson, V.W. (1992). Boston Naming Test: Shortened version for use in Alzheimer's disease. *Journal of Gerontology, 47*(3), 164–168. https://doi.org/10.1093/geronj/47.3.p154

Macizo, P. (2016). Phonological coactivation in the bilinguals' two languages: Evidence from the color naming task. *Bilingualism*, *19*(2), 361. https://doi.org/10.1017/S136672891500005X

McDaniel, M. A. (2006). Estimating state IQ: Measurement challenges and preliminary correlates. *Intelligence*, *34*(6), 607-619. http://dx.doi.org/10.1016/j.intell.2006.08.007

Mendoza, L., Garcia, P., Duara, R., Rosselli, M., Loewenstein, D., Greig-Custo, M. T., Barker, W., Dahlin, P., & Rodriguez, M. J. (2020). The effect of acculturation on cognitive performance among older Hispanics in the United States. *Applied Neuropsychology: Adult*, 1-9. https://doi.org/10.1080/23279095.2020.1725888

Menon, C., Hall, J., Hobson, V., Johnson, L., & O'Bryant, S. E. (2012). Normative performance on the executive clock drawing task in a multi-ethnic bilingual cohort: A project FRONTIER study. *International Journal of Geriatric Psychiatry*, *27*(9), 959-966. https://doi.org/10.1002/gps.2810

Meyers, J. E., & Meyers, K. R. (1995). Rey complex figure test under four different administration procedures. *The Clinical Neuropsychologist*, *9*(1), 63-67. https://doi.org/10.1080/13854049508402059

Michael, E. B., & Gollan, T. H. (2005). Being and becoming bilingual: Individual differences and consequences for language production. In J. F. Kroll & A. M. B. De Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 389–407). Oxford University Press.

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. https://doi.org/10.1007/s11065-008-9069-7

Moran, M. D. (2003). Arguments for rejecting the sequential Bonferroni in ecological studies. *Oikos*, *100*(2), 403-405. http://dx.doi.org/10.1034/j.1600-0706.2003.12010.x

Morrison, C., Kamal, F., & Taler, V. (2019). The influence of working memory performance on event-related potentials in young and older adults. *Cognitive Neuroscience*, *10*(3), 117–128. https://doi.org/10.1080/17588928.2019.1570104

Nakagawa, S. (2004). A farewell to Bonferroni: The problems of low statistical power and publication bias. *Behavioral Ecology*, *15*(6), 1044-1045. https://doi.org/10.1093/beheco/arh107

Nijdam-Jones, A., Rivera, D., Rosenfeld, B., & Arango-Lasprilla, J. C. (2019). The effect of literacy and culture on cognitive effort test performance: An examination of the test of memory malingering in Colombia. *Journal of Clinical and Experimental Neuropsychology*, *41*(10), 1015-1023. https://doi.org/10.1080/13803395.2019.1644294

Ong, G., Sewell, D. K., Weekes, B., McKague, M., & Abutalebi, J. (2017). A diffusion model approach to analysing the bilingual advantage for the Flanker task: The role of attentional control processes. *Journal of Neurolinguistics*, *43*, 28-38. https://doi.org/10.1016/j.jneuroling.2016.08.002

Paap, K. R., & Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cognitive psychology*, *66*(2), 232-258. http://dx.doi.org/10.1016/j.cogpsych.2012.12.002

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. https://doi.org/10.1016/j.cortex.2015.04.014

Pearson, N. C. S. (2009). *Advanced clinical solutions for WAIS-IV and WMS-IV: Administration and scoring manual*. The Psychological Corporation.

Perneger, T. V. (1998). What's wrong with Bonferroni adjustments. *BMJ*, *316*(7139), 1236-1238. https://doi.org/10.1136/bmj.316.7139.1236

Puente, A. E., Perez-Garcia, M., Lopez, R. V., Hidalgo-Ruzzante, N. A., & Fasfous, A. F. (2013). Neuropsychological assessment of culturally and educationally dissimilar individuals. In *Handbook of Multicultural Mental Health* (pp. 225-241). Elsevier Academic Press. https://doi.org/10.1016/B978-0-12-394420-7.00012-6

Rai, J. K., An, K. Y., Charles, J., Ali, S., & Erdodi, L. A. (2019). Introducing a forced choice recognition trial to the Rey complex figure test. *Psychology & Neuroscience, 12*(4), 451. https://doi.org/10.1037/pne0000175

Razani, J., Burciaga, J., Madore, M., & Wong, J. (2007). Effects of acculturation on tests of attention and information processing in an ethnically diverse group. *Archives of Clinical Neuropsychology*, *22*(3), 333-341. https://doi.org/10.1016/j.acn.2007.01.008

Reitan, R. M. (1958). Validity of the trail making test as an indicator of organic brain damage. *Perceptual and Motor Skills, 8*(3), 271-276. https://doi.org/10.2466%2Fpms.1958.8.3.271

Reitan, R. M., & Wolfson, D. (1985). *The Halstead–Reitan neuropsycholgical test battery: Therapy and clinical interpretation*. Neuropsychological Press.

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

Rivera, D., Perrin, P. B., Weiler, G., Ocampo-Barba, N., Aliaga, A., Rodríguez, W., Rodríguez-Agudelo, Y., Aguayo, A., Longoni, M., Trapp, S., Esenarro, & Arango-Lasprilla, J. C. (2015). Test of memory malingering (TOMM): Normative data for the Latin American Spanish speaking adult population. *NeuroRehabilitation*, *37*(4), 719-735. https://doi.org/10.3233/nre-151287

Roberts, P. M., Garcia, L. J., Desrochers, A., & Hernandez, D. (2002). English performance of proficient bilingual adults on the Boston naming test. *Aphasiology*, *16*(4-6), 635-645. https://doi.org/10.1080/02687030244000220

Rosselli, M., Ardila, A., Araujo, K., Weekes, V. A., Caracciolo, V., Padilla, M., & Ostrosky-Solí, F. (2000). Verbal fluency and repetition skills in healthy older Spanish-English bilinguals. *Applied Neuropsychology*, *7*(1), 17-24. https://doi.org/10.1207/s15324826an0701_3

Rouleau, I., Salmon, D. P., Butters, N., Kennedy, C., & McGuire, K. (1992). Quantitative and qualitative analyses of clock drawings in Alzheimer's and Huntington's disease. *Brain and Cognition*, *18*(1), 70-87. https://doi.org/10.1016/0278-2626(92)90112-y

Ryan, Camille. (2013). Language use in the United States: 2011. American Community Survey Reports, ACS-22. U.S. Census Bureau, Washington, DC.

Sakai, K. L. (2005). Language acquisition and brain development. *Science, 310*(5749), 815–819. http://dx.doi.org/10.1126/science.1113530

Salazar, X. F., Lu, P. H., Wen, J., & Boone, K. B. (2007). The use of effort tests in ethnic minorities and in non-English-speaking and English as a second language populations. In

K. B. Boone (Ed.), *Assessment of feigned cognitive impairment: A neuropsychological perspective* (p. 405–427). The Guilford Press.

Samuel, S., Roehr-Brackin, K., Pak, H., & Kim, H. (2018). Cultural effects rather than a bilingual advantage in cognition: A review and an empirical study. *Cognitive science*, *42*(7), 2313-2341. https://doi.org/10.1111/cogs.12672

Savitz, D. A., & Olshan, A. F. (1995). Multiple comparisons and related issues in the interpretation of epidemiologic data. *American Journal of Epidemiology*, *142*(9), 904. https://doi.org/10.1093/oxfordjournals.aje.a117737

Schretlen, D. J., Buffington, A. L., Meyer, S. M., & Pearlson, G. D. (2005). The use of word-reading to estimate "premorbid" ability in cognitive domains other than intelligence. *Journal of the International Neuropsychological Society: JINS*, *11*(6), 784. https://doi.org/10.1017/S1355617705050939

Schwartz, S. J., Benet-Martínez, V., Knight, G. P., Unger, J. B., Zamboanga, B. L., Des Rosiers, S. E., Stephens, D. P., Huang, S., & Szapocznik, J. (2014). Effects of language of assessment on the measurement of acculturation: Measurement equivalence and cultural frame switching. *Psychological Assessment*, *26*(1), 100. https://doi.org/10.1037/a0034717

Spitzer, R. L., Kroenke, K., Williams, J. B., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: the GAD-7. *Archives of Internal Medicine, 166*(10), 1092-1097. https://doi.org/10.1001/archinte.166.10.1092

Stålhammar, J., Hellström, P., Eckerström, C., & Wallin, A. (2020). Neuropsychological test performance among native and non-native Swedes: Second language effects. *Archives of Clinical Neuropsychology*. https://doi.org/10.1093/arclin/acaa043

Statistics Canada. (2011). Linguistic characteristics of Canadians. Retrieved from http://www12.statcan.gc.ca/census-recensement/2011/as-sa/98-314-x/98-314-x2011001-eng.pdf.

Strutt, A. M., Scott, B. M., Shrestha, S., & York, M. K. (2011). The Rey 15-item memory test and Spanish-speaking older adults. *The Clinical Neuropsychologist*, *25*(7), 1253-1265. https://doi.org/10.1080/13854046.2011.609839

Thirthalli, J., & Rajkumar, R. P. (2009). Statistical versus clinical significance in psychiatric research—An overview for beginners. *Asian Journal of Psychiatry*, *2*(2), 74. https://doi.org/10.1016/j.ajp.2009.04.004

Tombaugh, T. N. (1997). The test of memory malingering (TOMM): Normative data from cognitively intact and cognitively impaired individuals. *Psychological Assessment*, *9*(3), 260. https://doi.org/10.1037/1040-3590.9.3.260

Walker, A. J., Batchelor, J., & Shores, A. (2009). Effects of education and cultural background on performance on WAIS-III, WMS-III, WAIS-R and WMS-R measures: Systematic review. *Australian Psychologist*, *44*(4), 216-223. https://doi.org/10.1080/00050060902833469

Walker, A. J., Batchelor, J., Shores, E. A., & Jones, M. (2010). Effects of cultural background on WAIS-III and WMS-III performances after moderate-severe traumatic brain injury. *Australian Psychologist*, *45*(2), 112-122. https://doi.org/10.1080/00050060903428210

Wechsler, D. (1997). *WAIS-III administration and scoring manual*. The Psychological Association.

Wechsler, D. (2008). *Wechsler adult intelligence scale–Fourth Edition (WAIS–IV)*. NCS Pearson.

Wilkinson, G. S., & Robertson, G. J. (2006). Wide range achievement test (WRAT4). *Psychological Assessment Resources.*

Woumans, E. V. Y., Santens, P., Sieben, A., Versijpt, J. A. N., Stevens, M., & Duyck, W. (2015). Bilingualism delays clinical manifestation of Alzheimer's disease. *Bilingualism: Language and Cognition*, *18*(3), 568-574. https://doi.org/10.1017/S136672891400087X

Table 1
List of Tests Administered

Test	Subtest	Abbreviation	Norms	Reference	VM
Boston Diagnostic Aphasia Examination	Boston Naming Test – Short Form	BNT-15	-	Mack, Freed, Williams, & Henderson, 1992	HVM
	Complex Ideational Material	CIM_{BDAE}	Heaton	Goodglass & Kaplan, 1972	HVM
Clock Drawing Test	-	CDT	-	Rouleau et al., 1992	-
Delis-Kaplan Executive System	Color Naming	$COL_{D\text{-}KEFS}$	Manual	Delis et al., 2001	HVM
	Word Reading	$WOR_{D\text{-}KEFS}$	Manual	Delis et al., 2001	HVM
	Color-Word Interference	$INT_{D\text{-}KEFS}$	Manual	Delis et al., 2001	HVM
Generalized Anxiety Disorder	-	GAD-7	-	Spitzer, Kroenke, Williams & Löwe, 2006	-
Patient Health Questionnaire	-	PHQ-9	-	Kroenke, Spitzer & Williams, 2001	-
Rey Fifteen-Item Test	Free Recall	Rey-15 FR	-	Rey, 1941	LVM
	Recognition		-	Boone et al., 2002	LVM
Rey-Osterrieth Complex Figure Test	Copy	-	-	Meyers & Meyers, 1995	LVM
	3-minute Recall	IR_{RCFT}	Manual	Meyers & Meyers, 1995	LVM
	Delayed Recall	DR_{RCFT}	Manual	Meyers & Meyers, 1995	LVM
	Yes/No Recognition	REC_{RCFT}	Manual	Meyers & Meyers, 1995	LVM
	Forced Choice Recognition	FCR_{RCFT}	-	Rai et al., 2019	LVM
Rey Word Recognition Test	-	WRT	-	Rey, 1941	HVM
Wechsler Adult Intelligence Scale – Third Edition	Digit Span	$DS_{WAIS\text{-}III}$	Manual	Wechsler, 1997	HVM
	Digit-Symbol Coding	$CD_{WAIS\text{-}III}$	Manual	Wechsler, 1997	LVM
Wechsler Adult Intelligence Scale – Fourth Edition	Symbol Search	$SS_{WAIS\text{-}IV}$	Manual	Wechsler, 2008	LVM
Wide Range Achievement Test – Fourth Edition	Single Word Reading	$Reading_{WRAT\text{-}4}$	Manual	Wilkinson & Robertson, 2006	HVM
Test of Memory Malingering	Trial 1	TOMM-1	-	Tombaugh, 1997	LVM
Trail Making Test	Trails A	TMT-A	Heaton	Reitan & Wolfson, 1985	LVM
	Trails B	TMT-B	Heaton	Reitan & Wolfson, 1985	LVM
Verbal Fluency	Animals	-	Heaton	Lezak, Howieson, Loring & Fischer, 2004	HVM
	Letter	FAS	Heaton	Lezak et al., 2004	HVM
	Emotion Word	EWFT		Abeare et al., 2017	HVM
Visual Analog Scale		V-5	-	Erdodi et al., 2019	LVM
Word Choice Test	-	WCT	-	Pearson, 2009	HVM

Note. Heaton: Demographically adjusted T-scores based on norms by Heaton, Miller, Taylor and Grant (2004); VM: Verbal mediation; LVM: Low verbal mediation; HVM: High verbal mediation.

Table 2 *Sample Characteristics*

	English Proficiency									
	NSE		LEP							
Test	*M*	*SD*	*M*	*SD*	*t*	*p*	*d*	χ^2	*p*	Φ^2
Age	24.1	*8.5*	24.2	*2.9*	-0.09	.930	-	-	-	-
Education	15.0	*1.9*	16.3	*2.1*	-3.07	.003	.69	-	-	-
AOA $_{English}$	0.2	*1.0*	9.6	*3.5*	-16.4	<.001	-3.66	-	-	-
GAD-7	5.5	*3.5*	4.8	*3.4*	0.90	.369	0.20	-	-	-
PHQ-9	5.4	*3.5*	5.6	*3.3*	-0.30	.767	-0.07	-	-	-
% Female	85%		57.5%		-	-	-	7.38	.007	0.09
% RH	95%		92.5%		-	-	-	.213	.644	0.00

Note. AOA $_{English}$: Age of acquisition for English; GAD-7: General Anxiety Disorder; PHQ: Patient Health Questionnaire; RH: Right-handed; NSE: Native speakers of English: LEP: Limited English proficiency.

Table 3 *Performance on Neuropsychological Tests with HVM as a Function of English Proficiency*

	English Proficiency							
	NSE		LEP					
Test	*M*	*SD*	*M*	*SD*	*t*	*p*	*d*	σ_1^2 vs. σ_2^2
Animals	46.9	*10.9*	30.0	*9.9*	7.27	<.001	1.62	.757
BNT-15 *Accuracy*	13.9	*1.2*	6.5	*3.1*	14.02	<.001	3.14	.000
BNT-15 *T2C*	43.4	*27.8*	185.8	*57.2*	-14.18	<.001	-3.17	.000
CIM *BDAE*	44.2	*13.8*	16.3	*15.4*	8.51	<.001	1.92	.808
Color *D-KEFS*	10.1	*2.1*	7.2	*3.0*	5.02	<.001	1.12	.026
DS *WAIS-III*	10.0	*2.7*	9.7	*2.6*	0.63	.528	-	.580
LDF	6.6	*1.2*	6.2	*1.2*	1.67	.099	-	.964
LDB	4.9	*1.4*	5.3	*1.4*	-1.12	.267	-	.601
EWFT	13.3	*6.1*	8.1	*2.7*	5.03	<.001	1.13	.013
FAS	43.0	*9.9*	34.0	*5.9*	4.95	<.001	1.11	.007
Reading *WRAT-4*	102.7	*11.7*	85.4	*10.2*	7.05	<.001	1.58	.792
Rey WRT	10.8	*1.7*	10.1	*2.3*	1.48	.142	-	.028
Stroop *D-KEFS*	11.8	*2.0*	8.7	*3.6*	4.89	<.001	1.09	.000
WCT *Accuracy*	49.7	*0.7*	49.1	*1.5*	2.16	.034	0.48	.016
WCT *T2C*	72.4	*20.6*	95.8	*38.1*	-3.32	.001	-0.77	.029
Word *D-KEFS*	10.7	*2.4*	10.0	*2.3*	1.34	.184	-	.580

Note. Animals: Category fluency (Demographically adjusted T-scores; *M* = 50, *SD* = 10); BNT-15: Boston Naming Test – Short Form; Accuracy: Number of pictures correctly named within the time limit out of 15 (raw score); T2C: Total time-to-completion (seconds); CIM*BDAE*: The Complex Ideational Material subtest from the Boston Diagnostic Aphasia Battery (Demographically adjusted T-scores; *M* = 50, *SD* = 10); D-KEFS: Delis-Kaplan Executive Function System (age-corrected scaled scores; *M* = 10, *SD* = 3); DS *WAIS-III*: Digit Span subtest of the Wechsler Adult Intelligence Scale – Third Edition (age-corrected scaled scores; *M* = 10, *SD* = 3); LDF: Longest digit span forward (raw score); LDF: Longest digit span backward (raw score); EWFT: Emotion Word Fluency Test (raw scores); FAS: Letter fluency (Demographically adjusted T-scores; *M* = 50, *SD* = 10); Reading *WRAT-4*: Single word reading subtest from the Wide Range Achievement Test – Fourth Edition (Standard Scores; *M* = 100, *SD* = 15); Rey WRT: Rey Word Recognition Test (true positives); WCT: Word Choice Test; Accuracy: Number of words correctly recognized out of 50 (raw score); T2C: Total time-to-completion (seconds); HVM: High verbal mediation; NSE: Native speakers of English: LEP: Limited English proficiency; σ_1^2 vs. σ_2^2: *p*-value associated with Levene's test of homogeneity of variance.

Table 4

Performance on Neuropsychological Tests with LVM as a Function of English Proficiency

	English Proficiency							
	NSE		LEP					
Test	*M*	*SD*	*M*	*SD*	*t*	*p*	*d*	σ_1^2 vs. σ_2^2
CDT	9.7	*0.6*	8.4	*1.5*	4.98	<.001	1.12	<.001
CD $_{WAIS-III}$ ACSS	11.6	*2.4*	11.5	*2.6*	0.27	.791	-	.520
CD $_{WAIS-III}$ Recog	8.3	*1.1*	7.8	*1.2*	1.76	.082	-	.229
RFCT Copy	33.5	*2.5*	34.2	*1.7*	-1.39	.170	-	.058
RCFT IR	48.6	*14.8*	45.2	*12.4*	1.12	.268	-	.239
RCFT DR	46.9	*14.0*	44.2	*11.2*	0.95	.343	-	.086
RCFT REC	48.1	*12.9*	41.5	*14.3*	2.13	.037	0.48	.457
SS $_{WAIS-IV}$	11.3	*2.7*	11.0	*2.3*	0.58	.563	-	.599
TMT-A	41.0	*14.9*	37.9	*10.8*	1.07	.290	-	.169
TMT-B	47.8	*11.8*	42.9	*9.6*	2.04	.045	0.46	.211
TOMM-1	47.4	*3.1*	48.6	*2.0*	-2.14	.036	0.48	.081
Rey-15 FR	14.9	*0.5*	14.6	*1.1*	1.87	.065	-	.000

Note. CDT: Clock drawing test (raw score); CD $_{WAIS-III}$: Digit-Symbol Coding subtest of the Wechsler Adult Intelligence Scale – Third Edition (age-corrected scaled scores; *M* = 10, *SD* = 3); RCFT: Rey Complex Figure Test (T-scores; *M* = 50, *SD* = 10); IR: Immediate Recall; DR: Delayed Recall; REC: Yes/No recognition; SS $_{WAIS-IV}$: Symbol Search subtest of the Wechsler Adult Intelligence Scale – Fourth Edition (age-corrected scaled scores; *M* = 10, *SD* = 3); TMT: Trail Making Test (Demographically adjusted T-scores; *M* = 50, *SD* = 10); TOMM-1: Test of Memory Malingering – Trial 1 (raw score); Rey-15 FR: Rey Fifteen-Item Test (free recall raw scores); LVM: Low verbal mediation; NSE: Native speakers of English: LEP: Limited English proficiency; σ_1^2 vs. σ_2^2: *p*-value associated with Levene's test of homogeneity of variance.

Table 5

Performance on Neuropsychological Tests with HVM Compared to the Normative Mean as a Function of English Proficiency

	English Proficiency							
	NSE				LEP			
Test	M_{Diff}	*t*	*p*	*d*	M_{Diff}	*t*	*p*	*d*
Animals	-3.1	*-1.79*	.081	-	-20.0	-12.9	<.001	2.03
CIM *BDAE*	-5.8	*-2.66*	.011	0.42	-33.7	-13.7	<.001	2.20
Color *D-KEFS*	0.1	*0.22*	.825	-	-2.8	-6.02	<.001	0.95
DS *WAIS-III*	0.0	*0.06*	.954	-	-0.4	-0.87	.392	-
FAS	-7.0	*-4.45*	<.001	0.70	-16.0	-17.1	<.001	2.70
Reading *WRAT-4*	2.7	*1.47*	.150	-	-14.6	-9.08	<.001	1.44
Stroop *D-KEFS*	1.8	*5.85*	<.001	0.93	-1.3	-2.35	.024	0.37
Word *D-KEFS*	0.7	*1.85*	.073	-	0.0	0.00	1.000	-

Note. Animals: Category fluency (Demographically adjusted T-scores; *M* = 50, *SD* = 10); BNT-15: Boston Naming Test – Short Form; Accuracy: Number of pictures correctly named within the time limit out of 15 (raw score); T2C: Total time-to-completion (seconds); CIM_{BDAE}: The Complex Ideational Material subtest from the Boston Diagnostic Aphasia Battery (Demographically adjusted T-scores; *M* = 50, *SD* = 10); D-KEFS: Delis-Kaplan Executive Function System (age-corrected scaled scores; *M* = 10, *SD* = 3); $DS_{WAIS-III}$: Digit Span subtest of the Wechsler Adult Intelligence Scale – Third Edition (age-corrected scaled scores; *M* = 10, *SD* = 3); LDF: Longest digit span forward (*z*-score); LDF: Longest digit span backward (*z*-score); EWFT: Emotion Word Fluency Test (*z*-score); FAS: Letter fluency (Demographically adjusted T-scores; *M* = 50, *SD* = 10); $Reading_{WRAT-4}$: Single word reading subtest from the Wide Range Achievement Test – Fourth Edition (Standard Scores; *M* = 100, *SD* = 15); HVM: High verbal mediation; NSE: Native speakers of English: LEP: Limited English proficiency.

Table 6

Performance on Neuropsychological Tests with LVM Compared to the Normative Mean as a Function of English Proficiency

Test	English Proficiency							
	NSE				LEP			
	M_{Diff}	*t*	*p*	*d*	M_{Diff}	*t*	*p*	*d*
CD *WAIS-III*	1.6	*4.22*	<.001	0.67	1.5	*3.59*	<.001	0.57
RCFT IR	-1.4	*-0.61*	.546	-	-4.8	*-2.47*	.018	0.39
RCFT DR	-3.1	*-1.41*	.166	-	-5.8	*-3.30*	.002	0.52
RCFT REC	-2.0	*-0.96*	.345	-	-8.5	*-3.67*	<.001	0.60
SS *WAIS-IV*	1.3	*3.09*	.004	0.49	1.0	*2.65*	.012	0.42
TMT-A	-9.0	*-3.80*	<.001	0.60	-12.1	*-7.08*	<.001	1.12
TMT-B	-2.2	*-1.18*	.244	-	-7.1	*-4.67*	<.001	0.74

Note. CD *WAIS-III* : Digit-Symbol Coding subtest of the Wechsler Adult Intelligence Scale – Third Edition (age-corrected scaled scores; *M* = 10, *SD* = 3); RCFT: Rey Complex Figure Test (T-scores; *M* = 50, *SD* = 10); IR: Immediate Recall; DR: Delayed Recall; REC: Yes/No recognition; SS *WAIS-IV* : Symbol Search subtest of the Wechsler Adult Intelligence Scale – Fourth Edition (age-corrected scaled scores; *M* = 10, *SD* = 3); TMT: Trail Making Test (Demographically adjusted T-scores; *M* = 50, *SD* = 10); NSE: Native speakers of English: LEP: Limited English proficiency.

VITA AUCTORIS

NAME:	Maame Adwoa Brantuo
PLACE OF BIRTH:	Accra, Ghana
YEAR OF BIRTH:	1997
EDUCATION:	Ghana Christian International High School, Dodowa, Ghana, 2015
	Calvin College, B.Sc., Grand Rapids, MI, 2019