2022

An Examination of the Relationship between Validity and Memory Measures in Retired NFL Players

Huda Abu-Suwa

AN EXAMINATION OF THE RELATIONSHIP BETWEEN VALIDITY AND MEMORY MEASURES IN RETIRED NFL PLAYERS

by

Huda Abu-Suwa

A Dissertation Presented to the College of Psychology
of Nova Southeastern University
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy in Clinical Psychology

NOVA SOUTHEASTERN UNIVERSITY

2021

Dissertation Approval Sheet

This Dissertation was submitted by Huda Abu-Suwa under the direction of the Chairperson of the Dissertation committed listed below. It was submitted to the School of Psychology and approved in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Clinical Psychology at Nova Southeastern University.

Approved:

10/6/21
Date of Defense

Charles Golden, PhD, Chairperson

Lisa Lashley, PsyD

Christian DeLucia, PhD

Soledad Arguelles-Borge, PhD

10/26/21
Date of Final Approval

Charles Golden, PhD, Chairperson

Statement of Original Work

I declare the following:

I have read the Code of Student Conduct and Academic Responsibility as described in the Student Handbook of Nova Southeastern University. This dissertation represents my original work, except where I have acknowledged the ideas, words, or material of other authors.

Where another author's ideas have been presented in this dissertation, I have acknowledged the author's ideas by citing them in the required style.

Where another author's words have been presented in this dissertation, I have acknowledged the author's words by using appropriate quotation devices and citations in the required style.

No copyright material was included in this document which required permission from the author or publisher.

Huda Abu-Suwa
Name

10/26/21
Date

Acknowledgements

I would like to start by thanking my dissertation committee for your support and guidance throughout this process. Additionally, I would like to recognize all of the wonderful and amazing professors at the NSU College of Psychology and my practicum supervisors outside of NSU. You all have influenced and shaped how I am as a student, a future clinician, and as a person.

I would like to give a special note of appreciation to Dr. Golden. It was an honor to be your student and mentee. You work endlessly to support your students, and it doesn't go unnoticed. You pushed me to be the best student I can be and made me feel very prepared for internship and beyond. Your clinical supervision, class lectures, research support, and even your stories are all things that I will carry with me throughout my future career as a neuropsychologist.

Thank you to my classmates and friends who supported me throughout my graduate career. I truly feel grateful to be in a program that fosters collaboration and friendship, and I'm lucky to say that I now have life-long colleagues and friends.

Thank you to my parents for all of the sacrifices you've made to allow me to be where I am today. Thank you to my sister and my husband for always being there for me. The support of my family means everything to me, and I don't think I could have gotten through this program without you all.

TABLE OF CONTENTS

List of Tables

ABSTRACT

AN EXAMINATION OF THE RELATIONSHIP BETWEEN VALIDITY AND MEMORY MEASURES IN RETIRED NFL PLAYERS

by

Huda Abu-Suwa

Nova Southeastern University

Neuropsychologists have increasingly become involved in assessing sports-related concussions; however, an important concern is the validity of the evaluations. This study examined the relationship between Performance Validity Tests (PVTs) and memory measures in a comprehensive standardized battery administered to retired NFL players, with the purpose of exploring how predictive PVTs are for memory performance in this population.

Hierarchical multiple regression analyses were used to evaluate the relationship between four PVTs (TOMM, MSVT, RDS, and Word Choice) and six memory tasks (WMS-IV LM I and LM II, VPA I and VPA II, VR I and VR II). A regression analysis was conducted for each memory test, for a total of six regression analyses. For each model, years played in the NFL, as well as MMPI-2-RF RCd, RC2, and RC7 scales were entered into the first block and the four PVTs were entered into the second block. Each memory subtest was entered as a dependent variable.

Results yielded significant findings for each of the regression models, demonstrating that PVTs accounted for a significant amount of the variance of memory performance beyond the effects of emotional functioning, and years in the NFL. MSVT

FR was found to be a significant predictor for each of the memory scales. Reliable Digit Span was a significant predictor for immediate memory subtests. Word Choice was a significant predictor for VPA II, and TOMM was a significant predictor of VR I and II.

While the results demonstrated significant relationships between PVTs and memory performance, these relationships may be impacted by cognitive abilities, rather than true effort put forth on performance. This is particularly true for MSVT Free Recall, RDS, and the TOMM. Emotional functioning also appeared to impact memory performance. These results have important implications, including that PVTs may not be valid for individuals with severe cognitive impairment and that alternatives to validity testing may be necessary. Additionally, mood difficulties may exacerbate poor performance on neuropsychological testing. Overall, caution must be taken when evaluating performance on PVTs and cognitive tests in order to differentiate between genuine cognitive impairment, emotional distress, and suboptimal effort.

Chapter I: Statement of the Problem

Concussions, particularly sports-related concussions, are a significant medical concern that is being explored in existing research and highlighted in the popular media. Increasing evidence and attention has focused on multiple concussions and the reported long-term neurocognitive and psychological dysfunction that may accompany it. Neuropsychologists have increasingly become involved in assessing and treating athletes with sports-related concussions, however an important issue that arises in their assessment is the validity of the evaluation and the effort put forth by the athletes being evaluated.

This study examined the relationship between neuropsychological validity measures and memory measures in a comprehensive standardized battery administered to retired National Football League (NFL) players. Participants consisted of retired NFL players who underwent a day-long neuropsychological evaluation as part of the NFL's concussion settlement program. Individuals who did not complete the WMS-IV or all of the validity measures were excluded from the analyses.

Hierarchical multiple regression analyses were used to evaluate the relationship between four validity measures (TOMM, MSVT, Reliable Digit Span, and Word Choice) and memory performance across six memory tasks: WMS-IV Logical Memory I and II, Verbal Paired Associates I and II, Visual Reproductions I and II. Specifically, regression analyses were used to evaluate the predicative ability of the validity tests for memory performance. A regression analysis was conducted for each memory test, for a total of six regression analyses. For each regression model, years played in the NFL, as well as MMPI-2-RF Demoralization (RCd), Low Positive Emotions (RC2), and Dysfunctional

Negative Emotions (RC7) were entered into the first block to assess their contribution to the models. The four validity measures (TOMM, MSVT, Word Choice, and Reliable Digit Span) were entered into the second block. Each memory subtest was entered as a dependent variable.

The purpose of this study was to examine the relationship between validity measures and memory measures in the standardized battery utilized by the NFL. More specifically, this study examined the relationship between validity measures and memory performance in retired NFL players and explored how predictive validity measures are for neuropsychological performance in this population.

Chapter II: Review of the Literature

General Information Regarding Traumatic Brain Injury

Traumatic brain injury (TBI) is a significant public health concern, not only in the United States but across the world. An estimated 57 million people have been hospitalized with a TBI worldwide, with about 1-2 million occurring in the United States alone (Frost el al., 2013; Langlois, et al., 2006). This rate, however, is thought to be an underestimation of TBIs occurring in the United States, as TBIs treated in outpatient settings and military facilities, or those with undiagnosed or untreated TBIs are often not reported in US national data (Langlois et al., 2006). In the past, there was no clear indication of the definition of a traumatic brain injury. Traumatic brain injury can generally be defined as an injury to the brain resulting from an impact or the acceleration or deceleration of the brain (Lezak et al., 2012). The Center for Disease Control (CDC) defines TBI "as a disruption in the normal function of the brain that can be caused by a bump, blow, or jolt to the head, or penetrating head injury" (para. 1). A recent consensus of the definition of a TBI comes from the Demographics and Clinical Assessment Working Group of the International and Interagency Initiative toward Common Data Elements for Research on Traumatic Brain Injury and Psychological Health. They define TBI as "an alteration in brain function, or other evidence of brain pathology, caused by an external force" (Menon et al., 2010, p. 1637).

TBIs can be classified in a variety of ways. They can be considered open or closed head injuries, such that open head injuries include injuries in which the skull and dura are penetrated, and closed head injuries include injuries in which the skull remains intact and the brain is not exposed. TBIs are also commonly classified by severity, including mild, moderate, and severe TBIs. Furthermore, TBIs can be classified by pathoanatomic type, outcome and/or prognosis (Friedland & Hutchinson, 2013). Other common indicators include the presence or absence of loss of consciousness and post-trauma amnesia; however, these have been found to be less reliable measures of a TBI (Faul & Coronado, 2015).

Generally, males are twice as likely to sustain a TBI as compared to females. Young children (less than 4 years old) and older adolescents (15-19 years old) are at a higher risk for sustaining a TBI as compared to other age groups, however older adults also have high incidences of TBIs, especially in hospital settings (Langlois et al., 2006). Other risk factors for TBIs include low SES, minority racial status, low educational attainment, unemployment, history of psychiatric disorder, and alcohol consumption (Frost et al., 2013; Lezak et al., 2012). Individuals in the military or in professional sports teams are also at a higher risk for sustaining a TBI. According to the CDC, the main causes of TBIs include falls, car accidents, struck by or against events, assaults, and other or unknown causes. This may, however, not fully describe all types of TBI occurrences, as many forms of TBIs, such as those that occur by professional athletes, are often not accounted for by national consensus data (Langlois et al., 2006). As a result, sports-related head injuries serve as a unique subcategory of TBIs. Specific attention and research have therefore been dedicated to better understand the unique aspects of sports-related head injuries.

Sports-Related Head Injuries

Athletes, particularly professional athletes, is a population that is especially impacted by head injuries. The most common type of head injury experienced by athletes is a mild TBI, which is also commonly referred to as a concussion. It is estimated that about 1.6 to 3.8 million sports-related concussions occur annually in the U.S., however this is an underestimation, as about 50% of sport-related concussions go unreported (Clark & Guskiewicz, 2016). This occurs for many reasons, including a lack of understanding of concussion symptoms, beliefs that the injury is not serious, or not wanting to be removed from playing. Meier and colleagues found that athletes reported significantly fewer symptoms to athletic trainers using ImPACT testing as compared to self-reported symptoms collected during a confidential psychiatric interview. Additionally, athletes who were cleared to play continued to underreport symptoms 9 days post-concussion, especially psychiatric symptoms (Meier et al., 2015).

Our understanding of the signs and symptoms of concussions has evolved over the past few decades and increased attention and care has been made towards treatment. For example, it was commonly believed that a concussion must result in the loss of consciousness, however recent evidence suggests that loss of consciousnesses occurs in 10% or less of concussions. Additionally, concussions were previously believed to occur from injury directly to the head, however concussions may occur from an injury or impact to the head, face, neck, or elsewhere on the body (Dziemianowicz et al., 2012). Concussions may include a wide variety of symptoms, including physical, cognitive, and emotional symptoms. Such symptoms may include headaches, anterograde and/or retrograde amnesia, disorientation, dizziness, fogginess, fatigue, trouble sleeping, sensitivity to light and/or noise, visual disturbances, irritability, issues with balance, nausea, trouble concentrating, and emotional disturbances (Clark & Guskiewicz, 2016; Dziemianowicz et al., 2012). Among athletes, the most commonly reported symptoms include headaches (94.2%), followed by dizziness (75.6%), concentration difficulty (54.8%), confusion (45.0%), light sensitivity (36.0%), and nausea (31.4%) (Marar et al., 2012).

While concussions are associated with a wide variety of symptoms that may cause functional impairment, the symptoms are often transient in nature. Recovery time may vary depending on a variety of factors, however typical recovery time is usually between one to two weeks. The impairment seen in concussions is not only caused by the physical force or impact exerted on the brain, but also by the metabolic changes that occur as a result of the injury. Large changes in potassium and sodium levels occur extracellularly, in addition to a release of glutamate. This likely occurs due to sheering and straining of neurons (Clark & Guskiewicz, 2016). Additionally, a decrease in cerebral blood perfusion is often seen, creating an imbalance between glucose and blood perfusion in the brain. Furthermore, gender differences are also seen in sports concussion rates and recovery time. While males are more likely to experience head injuries in the general population, females are more likely to experience sport-related head injuries among athletes (Covassin, et al., 2018). Additionally, female athletes have been found to display more self-reported symptoms, higher neurocognitive impairment, and require longer recovery times as compared to male athletes (Covassin et al., 2018).

Of particular concern is the rate of concussions experienced by football players. Football has one of the highest rates of sports-related concussions, with a concussion occurring in the NFL about once every other game (Clark & Guskiewicz, 2016). Additionally, many football players experience repeated concussions, an issue that has gained much attention from the media in recent years and sparked the controversy surrounding chronic traumatic encephalopathy (CTE).

Chronic Traumatic Encephalopathy

Substantial attention has recently been placed on a disorder known as chronic traumatic encephalopathy, or CTE. CTE is a progressive neurodegenerative disorder caused by repetitive mild traumatic brain injuries and is characterized by global deposits of hyperphosphorylated tau (p-tau) as neurofibrillary tangles (McKee et al., 2013). While CTE is typically associated with repetitive head injuries, evidence also suggests CTE may develop as a result of a single moderate or severe TBI (VanItallie, 2019). CTE was originally described by Harrison Martland in 1928. He reported on clinical aspects of what he called ‘punch drunk syndrome’ in former boxers who experienced neurological deterioration (McKee et al., 2013; Asken et al., 2016). In 1937 Millspaugh changed the term to dementia pugilistica, however it was later recognized that the disorder was associated with other populations besides boxers, such as football players and military personal. The name was then changed to progressive traumatic encephalopathy and later to chronic traumatic encephalopathy, to reflect a more general term for the neurological decline (VanItallie, 2019).

CTE is associated with various emotional and behavioral disturbances, including irritability, impulsivity, aggression, paranoia, and depression, which typically develop 8-10 years after experiencing repetitive mild TBIs (McKee et al., 2013). Other symptoms include memory impairment, executive dysfunction, language/speech difficulties, attention impairment, gait difficulties, suicidality, and the development of dementia and parkinsonism (McKee et al., 2013; Omalu et al., 2011; Asken et al., 2016). Behavioral and emotional changes may precede cognitive changes; however, findings are currently mixed (Asken et al., 2016). In addition to cognitive and behavioral disturbances, CTE has been associated with shortened life expectancy and premature death. When examining former professional football players, McKee and colleagues found that in a sample of 35 CTE victims, the mean age at symptom onset was 54.1 (*SD*=14.1) years and the mean age of death was 67.1 (*SD*=16.6) years (McKee et al., 2013). In another sample of 80 CTE victims, the majority of whom were football players, the mean age of death was found to be 54 (*SD*=23; VanItallie, 2019).

CTE is often mistaken for other types of neurodegenerative disorders, such as Alzheimer's disease, Parkinson's disease, and amyotrophic lateral sclerosis (Maroon et al., 2015). CTE, however, is associated with distinct neuropathological changes that can be used to differentiate it from other neurodegenerative disorders, including widespread atrophy of the cerebral cortex, medial temporal lobe, diencephalon and mammillary bodies with enlarged ventricles (McKee et al., 2013; VanItallie, 2019). Other distinct findings include extensive p-tau neurofibrillary tangles found in the frontal and temporal lobes, limbic system, and brainstem, as well as degeneration of axons and white matter fiber bundles, thinning of the corpus callosum, and an absence of amyloid-B peptide deposits (McKee et al., 2013; VanItallie, 2019).

The center for the Study of Traumatic Encephalopathy at Boston University has a brain bank available, in which it preserves the brains and spinal cords of individuals that experienced repetitive mild TBIs. By utilizing this data, researchers at the center have identified four main stages of CTE (McKee et al., 2013). Omalu and colleagues also developed a system for classifying CTE, consisting of four histomorphology CTE phenotypes (Omalu et al., 2011; Gaetz, 2017). These classification systems provide inconsistent descriptions of the neuropathology of CTE and have raised criticism. In an effort to establish consistency in diagnosing CTE, the National Institute of Health (NIH) held a consensus conference and determined that CTE can only be diagnosed post-mortem at autopsy (Gaetz, 2017). They also indicated that “abnormal tau immunoreactivity in neurons and glia, in an irregular, focal, perivascular distribution and at the depths of cortical sulci, was required for the diagnosis of CTE” (Gaetz, 2017, p. 132).

Despite current research on CTE, our understanding of the prevalence and causation of CTE is still largely unknown (Maroon et al., 2015). Additionally, the proportion of athletes with sports-related head injuries that develop CTE is currently unknown, however current evidence suggests that those with repeated concussions or sub-concussive head injuries are at a higher risk for developing CTE (VanItallie, 2019). This risk increases as the number of years played increases and can vary depending on player position, with increased risk for those positions that are commonly exposed to a high level of physical impact (VanItallie, 2019). Another risk factor for developing CTE is age (Maroon et al., 2015). While older age has been associated with increased prevalence of CTE, the relationship between age and CTE is poorly understood, as older age is also associated with other neurodegenerative diseases, which may mediate the relationship, if a comorbid disorder exists or if an individual is misdiagnosed with CTE (Maroon et al., 2015). Smith and colleagues proposed a theory that CTE may be a combination of the neurological changes due to head injuries and normal age-related changes in the brain that occur naturally. Other research indicated that a history of repeated concussions may accelerate the normal aging process (Asken et al., 2016). Aging is also associated with a number of psychosocial stressors, which could be associated with late onset emotional disturbance and cognitive impairment (Asken et al., 2016). Other factors contributing to the development of CTE include sleep disturbances, a history of neurodevelopmental disorders, and alcohol/drug use (Asken et al., 2016).

In addition to the mentioned risk factors, CTE may have genetic risk factors, however further histological research is needed to improve our understanding. Apolipoprotein E (ApoE) is an allele that is a genetic risk factor for Alzheimer's disease (Maroon et al., 2015). There has been some evidence of the presence of ApoE being associated with worsened cognitive deficits following a severe head injury, however in a systemic review of CTE in contact sports, Maroon and colleagues did not find a significant association between ApoE carriers in those with CTE as compared to the general population (Maroon et al., 2015). In a study on the histomorphologic phenotypes of CTE, Omalu and colleagues found that 10 out of the 14 professional athletes examined were positive for CTE, and that out of the 10, seven had a known ApoE genotype (Omalu et al., 2011). Each of the participants had at least one E3 allele, with five of them having E3/E3 and two of them having E3/E4. Additionally, substance use may be a risk factor for CTE. Maroon and colleagues found that 20% of CTE cases had a history of substance abuse, which is substantially higher than the 7.7% reported for US adults over their lifetime (Maroon et al., 2015). The clinical presentation of CTE, however, may be distorted by substance abuse, as abused substances may lead to the development of neurodegenerative changes (Maroon et al., 2015).

Overall, CTE is an important clinical issue for professional athletes, particularly professional football players. Further research is needed to better understand the relationship between repetitive head injuries and the development of CTE, the clinical presentation and neuropathology of CTE, and other risk factors for developing CTE. An important and growing area of research that helps to elucidate our understanding of CTE and sports-related head injuries overall is neuroimaging research.

Neuroimaging Findings of Sports-Related Head Injuries

Neuroimaging findings play an important role in the assessment and treatment of sports-related concussions. Findings, however, have been somewhat inconsistent, such that a consistent injury pattern has yet to be found. For example, past CT studies on boxers failed to demonstrate damage or atrophy in those with concussions and were only found in severe head injuries, however recent studies have demonstrated neurological changes in those with sports-related concussions (Jordan et al., 1988; Murdaugh et al, 2018). Various factors may contribute to these inconsistent findings, such as the severity and location of injury, as well as the type of neuroimaging technique being utilized.

Common neuroimaging techniques utilized for assessing sports-related head injuries include MRI and CT. Oftentimes CT and MRI results of concussed patients appear normal, making it difficult to assess for neurological changes or deficits associated with a sports-related concussion. MRI has been shown to be more effective in detecting subtle neurological damage as compared to CT (Kelly et al., 1988), however MRI does not consistently identify these subtle changes, especially subcortically. Another neuroimaging technique often used to examine sports-related head injuries is Diffusion Tensor Imaging (DTI). DTI measures water molecule diffusion in various portions of the brain and can be used to measure white and gray matter damage after a concussion. DTI has been used to demonstrate atrophy in the brain following a concussion and is more sensitive in detecting structural damage as compared to other imaging techniques, however the utilization of DTI in assessing sports-related concussions is limited, as increased research is needed to determine its usefulness (Dziemianowicz et al., 2012).

Cubon and colleagues used DTI to study white matter skeleton in individuals with sports-related concussions. They evaluated white matter fiber tracts in college athletes with sports-related concussions without loss of consciousness and who experienced symptoms for at least one month. Tract-based spatial statistics was used to evaluate fractional anisotropy and mean diffusivity of the white matter, and results indicated increased diffusivity in several white matter fiber tracts in the left hemisphere, including the inferior and superior longitudinal and fronto-occipital fasciculi, the retrolenticular part of the internal capsule, and the posterior thalamic and acoustic radiations (Cubon et al., 2011). Cubon and colleagues concluded that mean diffusivity as measured by DTI may be sensitive enough to detect mild structural changes that occur in sports-related concussions. Other DTI studies have also demonstrated structural abnormalities in those with sports-related concussions. In a systematic review of DTI in sports-related concussions, Gardner and colleagues found that seven out of the eight DTI studies they examined reported structural changes associated with sports-related concussions. These included structural changes in various areas of the brain, including corpus callosum, hippocampus, dorsolateral prefrontal cortex, precuneus, primary visual cortex, midbrain, internal capsule, putamen, and temporal lobe (Gardner et al., 2012).

Positron emission tomography (PET) has been used to measure metabolic and functional changes that occur after a concussion. Coughlin and colleagues conducted a pilot study on neuroinflammation and brain atrophy in former NFL players, using PET. Nine former NFL players were utilized for the study, as well as nine age-matched healthy controls. Participants underwent a neuropsychological evaluation, genotyping testing, and PET scans. Results indicated a significant increase in DPA-713 binding to the translocator protein (TSPO), which is a biomarker of brain injury and repair, in former NFL players as compared to healthy matched controls (Coughlin et al., 2015). Increased binding was found in the supramarginal gyrus and right amygdala. Additionally, decreased brain volume was found in the right hippocampus, as well as differential performance on tests of verbal learning and memory, in former NFL players as compared to controls. The authors concluded that PET is a promising neuroimaging technique that can be used to evaluate sports-related head injuries and the relationship between brain injury, cognitive deficits, and biomarkers (Coughlin et al., 2015). While PET studies have demonstrated metabolic and functioning changes in concussions, the use of PET is limited, as it is time consuming, expensive, and exposes individuals to radiation. Additionally, underlying structural damage is often undetected, and as a result, PET is often paired with a structural imaging technique such as CT or MRI (Dziemianowicz et al., 2012).

Single Photon Emission Computerized Tomography (SPECT) measures regional blood flow and can be used to evaluate perfusion and activation in the brain. Its use for assessing sports-related concussions, however, is limited as not enough research has been conducted to demonstrate its utility specifically for sports-related injuries. Amen and colleagues evaluated long-term brain functioning in 100 active and retired NFL players using SPECT imaging. Participants underwent SPECT imaging and completed the MicroCog Assessment of Cognitive Functioning, CPT-II, and the Mild Cognitive Impairment Screen. Results demonstrated decreased regional cerebral blood perfusion globally in NFL players as compared to healthy controls, particularly in the prefrontal, temporal, parietal, and occipital lobes, as well as in the anterior and posterior cingulate gyrus and in the cerebellum (Amen et al., 2011). Additionally, NFL players were also found to display lower performance across all neuropsychological measures, except for measures of spatial processing and reaction time. Amen and colleagues concluded that playing professional football is associated with a high risk for brain damage and that SPECT can be used to provide important clinical information used to evaluate sports-related head injuries (Amen et al., 2011).

fMRI has been shown to be one of the most clinically useful neuroimaging techniques in assessing sports-related concussions and can observe functional changes in the brain after a concussion, both in the acute setting and months after injury. Additionally, fMRI is better able to demonstrate relationships between brain functioning and neuropsychological functioning, making it a useful tool in evaluating concussion deficits. Murdaugh and colleagues demonstrated the use of fMRI to evaluate longitudinal changes in resting state connectivity and white matter in adolescent football players with and without sports-related concussions (Murdaugh et al., 2018). Participants underwent resting state fMRI and DTI and completed ImPACT. Acute changes and changes 21 days later were evaluated. Murdaugh et al. (2018) found hyperconnectivity within the posterior regions of the brain and hypoconnectivity within the anterior regions in those with sports-related concussions as compared to the control group in the acute stage. Additionally, DTI results indicated varied diffusion in the concussion group along the corticospinal tract and the superior longitudinal fasciculus in the acute stage. No differences in imaging were found at the follow up stage. ImPACT results correlated with resting connectivity at both stages.

Overall, neuroimaging plays an important role in understanding the neurological underpinnings, clinical presentations, and long-term effects of sports-related head injuries. Various neuroimaging techniques may be utilized to evaluate sports-related head injuries, each with their own advantages and disadvantages for clinical use. Further research is needed to evaluate best practices for using neuroimaging to assess sports-related head injuries, as well as to establish consistent patterns of clinical findings.

Neuropsychological Assessment

While neuroimaging is used to evaluate neurological deficits following sports-related head injuries, neuropsychological evaluations are utilized to evaluate changes in cognitive, emotional, and behavioral functioning. In sports neuropsychology settings, a combination of computerized and paper/pencil neuropsychological tests or a “hybrid” neuropsychological testing approach is often used. It is recommended that the hybrid model be used in evaluations, as current evidence suggests that paper/pencil and computerized assessments do not measure the same domains of cognitive functioning and that each assessment provides unique information (Echemendia et al., 2020).

A wide variety of traditional paper/pencil neuropsychological tests are used to evaluate concussions. These typically include tests that measure cognitive domains impacted by concussions, including learning and memory, attention and concentration, processing speed, and executive functioning (Merritt et al., 2017). Examples of tests commonly used to assess concussions include the Brief Visuospatial Memory Test (BMVT); Trails Making Test, Hopkins Verbal Learning Test (HVLT); WAIS subtests, including Cancellation, Symbol Search, Digit Span, and Coding; Controlled Oral Word Association Test (COWAT); Stroop; and the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS) (Randolph et al., 2005; Merritt et al., 2017). A computerized test that is commonly used is the Immediate Post-Concussion Assessment Cognitive Test (ImPACT). ImPACT is a computerized measure of attention, working memory, processing speed, response variability, and nonverbal problem solving (Dziemianowicz et al., 2012). It is commonly used for both baseline and post-injury assessment and can be administered by nonmedical individuals, increasing its utility and availability in a variety of settings.

Sideline assessment tools are also often used to provide a quick assessment of cognition and physical status after an individual sustains a head injury. One such assessment is the Sports Concussion Assessment Tool (SCAT), which can be used as a sideline assessment tool and also a tool for patient education (Yengo-Kahn et al., 2016). The SCAT takes about 15-20 minutes to complete and produces a composite score (Dziemianowicz et al., 2012). Several versions of the SCAT have been developed, with the most recent version being the SCAT-5. While the SCAT is widely used, caution should be taken, as players may underreport symptoms so as to not be pulled from a game. Another commonly used sideline assessment tool is the King-Devick Test (K-D). The K-D is a number naming test that measures eye movements and saccades, attention, and language (Dziemianowicz et al., 2012). It targets functions of the brainstem, cerebellum, and cerebral cortex, which may be impaired in patients after a concussion (Dziemianowicz et al., 2012). It can be administered in 1-2 minutes, making it very practical, however it has not been as thoroughly researched as other sideline assessments. Additionally, there may be a learning effect associated with the test (Dziemianowicz et al., 2012).

Overall, a wide variety of neuropsychological tests are utilized to evaluate cognitive, emotional, and behavioral changes following sports-related head injuries. These assessments may be administered in paper/pencil or computerized format, and oftentimes both test formats are utilized. Neuropsychological assessments may be used to establish a baseline level of cognitive functioning or may be used post-injury at sideline or in medical settings.

Memory and TBI

Memory is the most widely studied cognitive domain in TBI patients and substantial evidence exists to suggest that memory and learning can be impacted by TBIs (Vakil, 2005). Memory deficits are one of the most commonly reported forms of cognitive impairment after a sports-related concussion and can last for days after a concussion (Randolph et al., 2005). Furthermore, memory impairment is commonly implicated in TBIs generally, across TBI severity type and in repeated TBIs or CTE. Both verbal and visual memory are often impacted, as well as learning and retention. Additionally, prospective memory, or the ability to remember to do a previously planned action at a certain time or situation, has been found to be impaired in those with TBIs (Vakil, 2005). Furthermore, TBI patients demonstrate a reduced ability to utilize semantic information to aid memory and learning and display the same rate of forgetting of semantically organized information as compared to unrelated information. In addition, TBI patients have been shown to display slower learning rates and require more trials to encode and learn information as compared to controls (Vakil, 2005). Long-term memory and previously learned skills are often intact, however the ability to learn new skills may be impacted. In TBIs, memory has been found to recover more slowly than other cognitive functions, with deficits apparent years after injury, particularly in severe TBI patients (Vakil, 2005).

Several studies have demonstrated the impact of TBIs on memory functioning. Carlozzi and colleagues examined memory functioning in individuals with TBIs using the WMS-IV. One hundred individuals with either complicated mild/moderate or severe TBI were utilized, along with 100 matched controls from the WMS-IV normative dataset. Carlozzi et al. (2013) found that severe TBI participants had poorer performance across all WMS-IV indices and subtests as compared to the control group. Individuals with complicated mild/moderate TBI performed more poorly on all indices and on visual memory and visual working memory subtests. Furthermore, memory has been demonstrated to be impacted by TBIs even after controlling for practice effects. Lubrini et al. (2020) conducted a study in which they examined whether improvements were shown in neuropsychological performance in TBI patients after controlling for practice effect. Participants consisted of mild, moderate, and severe TBI patients, and all participants completed the Trails Making Test; Stroop Test; WAIS Digit Symbol-Coding, Symbol Search, Digits Forward and Digits Backwards; Verbal Fluency Test; and Logical Memory immediate recall and recognition. Patients completed the assessments during the acute phase of their TBI and approximately six months later. Results initially demonstrated that patients showed improvements across all tests, however after controlling for practice effect, improvements were seen on only on the Trails B, Digit Symbol-Coding, Symbol Search, Stroop Color-Word and Digits Backwards, suggesting that cognitive deficits were present (Lubrini et al., 2020).

Memory deficits have repeatedly been found in retired athletes with a history of sports-related concussions. Specific findings include deficits in episodic memory, short-term memory, visual memory, and verbal memory, as well as increased retroactive interference on learning tasks (Asken et al., 2016; Coughlin et al., 2015; VanItallie, 2019). Additionally, research has demonstrated that NFL players may experience decreased cognitive reserve, or the ability to retain premorbid functioning after experiencing a traumatic event. Randolph and colleagues (2013) demonstrated that retired NFL players had symptoms related to mild cognitive impairment (MCI), such that they exhibited deficits across multiple cognitive domains, including memory and decreased cognitive reserve, as compared to normal matched controls. Furthermore, retired NFL players with a history of multiple concussions have been found to be five times more likely to develop MCI and three times more likely to have subjective memory complaints (Randolph et al., 2013).

Memory is the most common cognitive deficit experienced by those with sports-related head injuries and the mostly widely studied cognitive domain in those with TBIs. TBIs can impact various forms of memory, including short-term, delayed, verbal, visual, episodic, and prospective memory. TBIs have also been linked to increased interference, decreased cognitive reserve and semantic utilization, and slower learning rates. Retired athletes have been shown to display memory impairment and display a substantial risk for memory deficits and developing MCI.

Effort and Validity

Effort is an import construct in neuropsychological testing and one that has received increased attention in recent years. Effort may be impacted by a variety of factors, including physical, psychological, and/or emotional state; lack of sleep; hunger; or diminished motivation or interest in an assessment. Additionally, individuals may not put forth full effort for the purpose of intentionally feigning performance for an external or instrumental gain, which is known as malingering. While malingering and suboptimal effort are often considered to be synonymous, suboptimal effort does not equate to malingering and significant consideration must be given to determine if an individual is malingering. One method of assessing for malingering is by using the Slick criteria (Slick, Sherman, & Iverson, 1999) which considers information from psychometric, behavioral, and external sources to determine if an individual may exhibit possible, probable, or definite malingering.

Effort can be assessed using performance validity tests (PVTs). Performance validity tests are used by neuropsychologists to determine whether an individual's performance in an evaluation is valid or invalid, in which valid results reflect results of true neurocognitive functioning and invalid results reflect results that are highly impacted by an individual's effort or engagement in testing (Greher & Wodushek, 2017). PVTs can be stand-alone tests designed specifically for assessing effort, or it may be specific embedded scores within traditional neuropsychological tests that have demonstrated ability in assessing effort (Greher & Wodushek, 2017). Examples of embedded validity test include the Reliable Digit Span, California Verbal Learning Test (CVLT-2) forced choice task, the Wisconsin Card Sorting Test (WCST) failure to maintain set score, and the Weschler Memory Scale (WMS-IV) Logical Memory II recognition and Visual Reproductions II recognition tasks (Greher & Wodushek, 2017; Kirkwood et al., 2011). Commonly used stand-alone tests of effort include the Test of Memory Malingering (TOMM), Word Memory Test (WMT), Medical Symptom Validity Test (MSVT), Victoria Symptom Validity Test, Dot Counting Test, and the Rey 15-item Test (Greher & Wodushek, 2017). PVTs are used not only in settings in which an examinee's performance may be questionable, such as in forensic cases or cases involving a secondary gain, but also in general clinical settings, as clinical judgement is often insufficient to accurately determine test validity without these more objective measures (Greher & Wodushek, 2017).

Generally, PVTs are designed to be resistant to the effects of cognitive impairment, however recent evidence suggests that some PVTs may be impacted by memory deficits, particularly if the PVT relies on latent memory variables. For example, the MSVT is generally found to be a strong PVT and assessment of feigned cognitive impairment (Armistead-Jehle & Hansen, 2016; Green et al., 2011). The Free Recall subtest, however, has been shown to load more on to memory than effort when examining factor analyses of the MSVT and memory tests (Armistead-Jehle & Hansen, 2016). This pattern was found when examining both the traditional MSVT and nonverbal version of the test (NV-MSVT). Other studies examining the Word Memory Test, the test from which the MSVT was derived, have demonstrated that the Free Recall and Paired Associates subtests functioned as measures of episodic memory (Armistead-Jehle et al., 2016). Furthermore, Willis and colleagues found that individuals failed the WMT Immediate and Delayed Recall subtests and that failures were related to cognitive deficits, including deficits in memory and attention (Willis et al., 2011).

The Reliable Digit Span (RDS) is one of the most widely used embedded validity test. It has been shown to demonstrate strong specificity and moderate sensitivity (Jasinski et al., 2011). While the common cutoff score utilized for RDS is 7, this score has been demonstrated to be inappropriate for some populations, including those with strokes, intellectual disability, memory disorders, and those who are non-English native speakers, causing a higher false positive rate of suboptimal effort for these populations (Maiman et al., 2019). Furthermore, RDS performance has been found to be associated with performance on tasks of memory, attention, and processing speed, as well as overall IQ, suggesting that it is not an independent measure of effort but that it is also linked to cognitive functioning (Maiman et al., 2019).

The TOMM has been shown to be a very strong PVT and resistant to the effects of various cognitive, psychiatric, and medical disorders. Initial validation studies demonstrated that the TOMM was effective at predicting sub-optimal effort even in those with severe memory impairment, however recent evidence suggests that the TOMM may not be effective for those with dementia, particularly when the standard cutoff score of 45 is utilized (Teichner & Wagner, 2004). Furthermore, high false positive rates have been identified for those with dementia even when using cutoff scores of 42 or 40 (Teichner & Wagner, 2004). TOMM performance has been found to be impacted by level of education and literacy, as well as severe memory impairment such as amnesia, increasing the risk of false positive results (Oudman et al., 2019; Nijdam-Jones et al., 2019).

As discussed previously, underreporting symptoms is a fairly common occurrence among athletes. Athletes may underreport symptoms for a variety of reasons, most importantly so as to not get pulled from a game or series of games. As such, it is not surprising that athletes may also feign poor performance. In an interview with NFL player Peyton Manning, Manning indicated that he intentionally performed poorly on baseline concussion tests so as to not be pulled from a game after getting injured (Erdal, 2012). Additionally, athletes may attempt to perform poorly on post-injury cognitive testing. Reasons for poor performance on post-injury cognitive testing may be related to a secondary gain or may be a result of emotional distress related to the injury and/or other factors. Athletes may attempt to do poorly on post-injury testing for fear of re-injury or losing academic accommodations, or they may feel pressure from others to remain in the sick role (Chase et al., 2018). Additionally, sub-optimal effort at post-injury testing may be related to internal and/or external pressure to immediately return to premorbid levels of functioning or uncertainty about one's identity after he/she is no longer concussed (Chase et al., 2018). Whether individuals can successfully perform poorly on neuropsychological testing without reaching thresholds on validity tests is a question that has been debated and researched. In a study on coached feigning, Jelicic and colleagues found that even when participants were instructed to perform poorly and had previous experience with tests, "sandbagging" is difficult to achieve without being identified (Jelicic et al., 2011).

In conclusion, effort is an important factor to consider when conducting neuropsychological evaluations in general clinical, forensic, and sports medicine settings. Effort may be impacted by physical, psychological, and/or emotional state, and may also be related to diminished motivation or interest in an assessment. Effort tests may be administered as a stand-alone test or as an embedded test within a commonly used neuropsychological test. While poor performance on effort tests is commonly equated with malingering, suboptimal effort does not equate to malingering and significant consideration must be given to determine if an individual is malingering. Additionally, careful consideration must be given when utilizing effort tests for those with cognitive impairment, as recent evidence suggests that effort performance may be impacted by cognitive impairment.

Emotional Functioning and Neuropsychological Performance

As indicated above, an individual's emotional state may impact the level of motivation and effort put forth during a neuropsychological evaluation, which in turn may impact neuropsychological performance. More specifically, anxiety and depression have commonly been found to impact neuropsychological testing. This may be due to symptoms related to anxiety and depression, such as psychomotor retardation, fatigue, reduced motivation, and general cognitive inefficiency (Gass, 1991; Lezak et al., 2012). Additionally, negative expectations about one's capabilities may impact performance. High levels of anxiety and depression have been associated with slowed thinking and responding, mental blocks, distractibility and difficulties with memory (Lezak et al., 2012). While anxiety and depression may impact performance during testing, many studies have found that anxiety and depression do not substantially impact neuropsychological results in TBI patients, psychiatric patients, or those with comorbid disorders (Lezak et al., 2012). For example, in a study examining the effects of depression and anxiety on neuropsychological performance in those receiving pre-surgical evaluations, Tsushima and colleagues found age and education to be significant predictors of neuropsychological performance, however depression and anxiety were non-significant (Tsushima et al., 2005). In a similar study evaluating the impact of depression on neurocognitive performance, Rohling and colleagues found that depression did not impact objective neurocognitive functioning, however depressed patients reported more cognitive problems on self-report measures (Rohling et al., 2002). Other studies, however, have noted cognitive differences between those with and without depression. Mohn and Rund (2016) found those with depression to perform significantly worse on measures of working memory, processing speed, attention, learning, and problem solving as compared to those without depression.

While many studies suggest that anxiety and depression do not impact neuropsychological performance, many patients with emotional difficulties report substantial cognitive deficits, including difficulties with memory, attention, processing speed, and executive functioning (Perini et al., 2019). These difficulties have been found to persist even in the remission phases of these disorders (Perini et al., 2019). Pseudodementia is a term commonly used for individuals displaying cognitive deficits in the context of emotional dysfunction, most commonly depression. Pseudodementia is commonly seen in older adults and increases the chances of developing dementia (Perini et al., 2019). While pseudodementia can be difficult to differentiate from dementia, key differences have been noted between the disorders in terms of clinical presentation, onset, progression, patient insight, and cognitive performance (Perini et al., 2019; Kang et al., 2014). As a result, careful consideration should be given when evaluating patients presenting with emotional dysfunction, cognitive difficulties, and comorbid conditions, such that neurological and neuropsychological clinical patterns must be identified for accurate differentiation (Perini et al., 2019; Lezak et al., 2012; Kang et al., 2014).

Substantial evidence exists suggesting that retired professional athletes experience emotional difficulties, most notably depression. In a neuroimaging study on retired NFL players, Amen and colleagues (2011) found that retired NFL players had significantly higher rates of depression as compared to the general population and displayed neuroimaging findings that have been associated with resistant depression and suicide (Amen et al., 2011). Additionally, Didehbani and colleagues (2013) conducted a study examining depression symptomatology in retired NFL players. They found that retired NFL players endorsed more cognitive, affective, and somatic symptoms of depression as compared to matched controls (Didehbani et al., 2013).

The causes of emotional difficulties in retired NFL players are multifaceted and may be related to premorbid factors, history of concussions and/or substance use (Gaetz, 2017; Didehbani et al., 2013). Additionally, difficulties that arise after retirement, such as loss of identity, loss of social support, low sense of belongingness, difficulty adjusting to life after football, and the development of chronic pain have also been associated with increased emotional distress (Gaetz, 2017).

In conclusion, emotional difficulties may impact an individual’s ability to engage in neuropsychological testing and may cause difficulties with memory, attention, fatigue, mental blocking, reduced motivation, and slowed thinking and responding. Mixed results have been found regarding whether anxiety and depression may impact neuropsychological performance. Nevertheless, individuals with anxiety and depression often endorse cognitive deficits, such as difficulties with memory, attention, processing speed, and executive functioning. Pseudodementia is a term used to describe the cognitive deficits associated with emotional dysfunction. Pseudodementia often mimics dementia, however neurological and neuropsychological findings can be utilized to differentiate the disorders. Emotional dysfunction serves as an important factor to consider when evaluating retired NFL players, as these individuals often present with depression and anxiety.

Clinical Relevance

The controversy surrounding the NFL and sports-related concussions has received significant attention from the media and general population in recent years. As such, the NFL and other sports organizations have developed an increasing interest in properly managing and addressing sports-related concussions. While these efforts are currently being made, they do not address the needs of former NFL players, as many of these players have retired prior to the implementation of these efforts. As a result, retired NFL players possess unique qualities and challenges from those who are currently playing in the NFL.

The negative impacts of repeated head injuries, particularly in retired football players, have been established in the literature, and include substantial cognitive decline, most notably in memory. While neuropsychologists are becoming increasingly involved in evaluating and treating professional athletes, their involvement is highly dependent on the validity of their evaluations, which may be compromised if a given athlete is not putting forth adequate effort, whether intentional, due to emotional difficulties, or some other factor. This poses a challenge for neuropsychologists, as the frequency with which individuals have malingered cognitive symptoms has increased over time, particularly as related to personal injury, disability, criminal, or medical cases (Delain et al., 2003; Mittenberg et al., 2002). Additionally, over half of retired NFL players have reported hiding concussions sustained during a game from training staff (Kerr et al., 2018). By conducting this study, researchers can provide more information regarding the relationship between validity measures and memory performance and help clinicians make better decisions and impressions regarding client functioning and treatment.

Purpose

The purpose of this study was to examine the relationship between validity measures and memory measures in the NFL battery. More specifically, this study examined the relationship between validity and neuropsychological performance in retired NFL players and explored how predictive validity measures function for memory performance in this population.

Research regarding long-term outcomes of repeated concussions from football has been well established, however no research has examined the validity of the currently used battery in retired players. Currently, some research exists examining validity testing in athletes, however much of the focus of this literature is on underreporting symptom versus on neuropsychological performance. Additionally, much of the research focuses on high school and college level athletes, and few studies examine professional athletes, such as NFL players. As such, this study attempted to provide relevant clinical information regarding the assessment of a unique population, retired NFL players. The purpose of examining the relationship between the validity measures and memory measures in this battery was to evaluate the extent to which validity measures are associated with and/or predict memory performance. This will provide important information regarding decisions made by the NFL concussion settlement, whether these validity measures provide valid information regarding neuropsychological performance, and what it means to meet or not meet criteria for cognitive impairment. Finally, this study addressed the lapse in information regarding the use of these various validity measures in providing neuropsychological assessments for retired NFL players.

Hypothesis

It was hypothesized that the four validity measures (TOMM, MSVT, RDS, and Word Choice) would significantly predict and account for performance on WMS-IV memory subtests when controlling for emotional functioning and years played in the NFL.

Justification

Evidence exists to suggest that effort and validity measures are a significant predictor of performance across various neuropsychological tests and domains, including attention, executive functioning, memory, and intelligence (Lindem et al., 2003; Perna & Loughan, 2014). These findings are not only seen in healthy samples, but also in brain injured individuals. Green (2007) found that in a TBI sample, performance of memory and learning were influenced more by effort measures than the effects of brain injury, with effort accounting for 34% of the variance. Additionally, Armistead-Jehle and Hansen (2016) conducted a study in which they found that MSVT variables accounted for 7-33% of the variance of scores on the RBANS immediate and delayed memory subtests. Moreover, when comparing validity measure performance and memory performance, those who put forth less than 50% effort on validity tasks have been found to perform about two standard deviations below those who put forth full effort on verbal memory tasks and about one standard deviation below on visual memory tasks (Green, 2007). Additionally, scores were found to decrease as effort performance decreased. These findings suggested a strong relationship between memory performance and effort measures (Green, 2007).

Limited research has been conducted examining WMS-IV memory performance and the variance explained by effort. Additionally, in the few studies that have examined the variance of memory performance explained by effort, only one effort measure was used to evaluate the relationship between effort and memory. It is hypothesized that using multiple effort measures will provide an incrementally increased explanation of variance for the WMS-IV subtests and result in a higher amount of variance explained by the effort measures than is suggested in the current literature. Furthermore, considering that many of the effort measures used in the current study have been found to tap into underlying memory constructs, including Reliable Digit Span, and MSVT, performance on these tasks are likely highly related to memory performance and will likely demonstrate a strong relationship with WMS-IV subtests (Armistead-Jehle & Hansen, 2016; Miller et al., 2011; Maiman et al., 2019).

While emotional functioning, specifically anxiety and depression, have been found to cause slowed thinking, distractibility, reduced motivation, fatigue, and psychomotor retardation during neuropsychological testing, many studies have found that anxiety and depression do not substantially impact neuropsychological results in TBI patients, psychiatric patients, or those with comorbid disorders (Lezak et al., 2012; Rohling et al., 2002). As a result, it is hypothesized that the validity measures will account for a significant amount of the variance of the six WMS-IV subtests, above and beyond the effects of anxiety and depression.

Additionally, it is hypothesized that that the validity measures will account for a significant amount of the variance of WMS-IV subtests, above and beyond the effects of years played in the NFL. Findings regarding whether years played in the NFL are related to cognitive performance have been inconsistent. Stamm and colleagues (2015) examined the effects of age of exposure to football by comparing neuropsychological performance in retired NFL players who were exposed to tackle football before age 12 and after age 12. They found that retired NFL players who were exposed to football before the age of 12 performed significantly worse than those who were exposed after age 12 across cognitive domains (Stamm et al., 2015). In contrast, Fields and colleagues (2020) examined the relationship between years played in the NFL and cognitive outcomes in retired NFL players, and found that years played in the NFL was not associated with cognitive performance across various domains (Fields et al., 2020). Studies linking years played in the NFL to subsequent cognitive impairment have suggested that this link is related to increased risk of sustained concussions (Stamm et al., 2015; Maroon et al., 2015). As mentioned above, memory performance has been found to be influenced more by validity measures than the effects of brain injury. As such, even when accounting for years played in the NFL and potential increased exposure to sports-related head injuries, it was hypothesized that validity measures would account for a significant amount of the variance of the six WMS-IV subtests.

Chapter III: Method

Participants

Participants were selected from an archived, de-identified database of adult retired NFL players (*N*=126). Participants were referred for testing to a private licensed neuropsychologist as part of their involvement with the NFL concussion settlement. Individuals were either self-referred or referred by their attorney to see this neuropsychologist. Participants were included in the study if they completed a comprehensive neuropsychological evaluation and had all scores available for the WMS-IV, MMPI-2-RF, TOMM, MSVT, Reliable Digit Span, and ACS Forced Word Choice.

Participants included in the study were between the ages of 31 and 83, with a mean age of 48.97 years (*SD*=10.42). The majority of the participants had a bachelor's degree (56.5%) and played professionally for an average of 6.11 years (*SD*=3.54). The majority of the participants identified as African American (73.2%) and 26.8% identified as Caucasian. All of the participants were male. The sample included a broad range of player positions and were comprised of the following: Defensive Back=20.2%; Wide Receiver=16.9%; Defensive Lineman=14.5%; Running Back=13.7%; Offensive Lineman=13.7%; Linebacker=11.3%; Tight End=4%; Quarterback=3.2%; Special Teams=1.6%; and Fullback= .8%. English was the primary language of all participants and was the language used in testing. Information regarding medical and/or psychiatric diagnoses were not included in this study. Test of Premorbid Functioning (TOPF) scores ranged from a standard score of 87 to a standard score of 118, low average intelligence to high average intelligence. The mean premorbid estimate of intellectual functioning is 103.12 (SD=5.94).

Measures

Weschler Memory Scale- Fourth Edition (WMS-IV)

The Weschler Memory Scale-Fourth Edition (WMS-IV) is a measure of visual and verbal memory. The WMS-IV assesses immediate, delayed, and working memory functioning. The WMS-IV contains an Adult battery for those ages 16-69 and an Older Adult battery for those ages 65-90. The Adult battery consists of 10 subtests that produce five index scores, namely an Auditory Memory, Visual Memory, Visual Working Memory, Immediate Memory, and Delayed Memory Index score. The Older Adult Battery consists of six subtests that produce four index scores. These include the same index scores as the Adult battery scores, excluding the Visual Working Memory index. Participants in the study completed the following subtests of the WMS-IV: Logical Memory I, II, and Recognition; Verbal Paired Associated I, II, and Recognition; Visual Reproductions I, II, and Recognition. For the purpose of this study, age-corrected scaled scores of Logical Memory I and II, Verbal Paired Associated I and II, and Visual Reproductions I and II were utilized.

Verbal memory, visual memory, learning, and prospective memory have all been found to be impacted by TBI, however in those with mild or moderate TBI, visual memory has been found to be more impacted than auditory memory (Vakil, 2005; Carlozzi et al., 2013). Additionally, individuals with complicated mild/moderate TBI have been found to perform poorly on WMS-IV subtests of visual memory and visual working memory as compared to controls, but not on auditory memory tasks (Carlozzi, et al., 2013).

The normative sample from the WMS-IV consisted of 1,400 people, 900 of whom were utilized for the Adult Battery and 500 of whom were utilized for the Older Adult battery. The sample was divided into 14 age bands ranging from 16-90, and 100 individuals were included in each age band. Race, gender, and education demographics of the sample were proportional to the 2005 census.

Test of Memory Malingering (TOMM)

The Test of Memory Malingering (TOMM) is a well-known measure of effort and is often used to assess for malingering. The TOMM focuses on memory and is used to discriminate memory-impaired individuals from those who are feigning (Tombaugh, 1996). The test is comprised of 50-line drawings of common objects that are shown to an individual for three seconds, and one-second intervals. The test includes three trials: two learning trials and one retention trial. During both learning trials, the individual is shown the 50 drawings and then is shown 50 recognition panels, one at a time. The panels each contain two pictures: one of a previously presented target picture and one new picture. The individual is asked to choose which picture is the picture that was seen before. During the retention trial, the individual is only shown the 50 recognition panels and asked to choose which picture is the correct one. A criterion score of 45 out of 50 (90% correct) is required for each trial, such that a raw score less than 45 indicates a lack of effort and/or malingering (Tombaugh, 1996). As such, raw scores will be utilized for this study.

The TOMM has strong face validity as a test of memory and recognition, making it difficult to recognize as a test of effort and/or malingering (Tombaugh, 1997). Researchers have suggested, however, that even if an individual is informed of or coached on the nature of the test, clinicians can still discern those who are malingering from those who are not (Jelicic et al. 2011). The TOMM was normed using samples of individuals with severe memory impairments, as well as those with general cognitive deficits (Tombaugh, 1996). The participants consisted of 475 healthy controls and 161 cognitively impaired individuals. The cognitive impairment sample was comprised of those with conditions such as traumatic brain injuries, aphasia, Alzheimer's dementia, vascular dementia, amnesia, ADHD, learning disabilities, stroke, and Parkinson's disease. The TOMM has since been validated on other sample populations, including in children and those with psychotic disorders (Donders, 2005; Duncan, 2005).

Medical Symptom Validity Test (MSVT)

The Medical Symptom Validity Test (MSVT) is a computerized verbal memory test used to assess effort and malingering (Green, 2004). The MSVT is based on Green's Word Memory Test (WMT) and is a short version of the WMT. The MSVT consists of 10-word pairs that are semantically related and contains four subtests: Immediate Recognition, Delayed Recognition, Paired Associates, and Free Recall subtests. The individual is presented with the word pairs two times and then completes the Immediate Recognition subtest, in which the individual is asked to choose the target words from 20 new word pairs. After a ten-minute delay, the Delayed Recognition subtest is administered, which is similar to the Immediate Recognition task except that new distractor words are used. During the Paired Associates subtest, the individual is given the first word from each word pair and is asked to provide the second word. During the Free Recall subtest, the individual is asked to recall as many words as possible from the original list. For the purpose of this study, percentile scores were utilized for analyses.

The MSVT has high specificity and sensitivity to detecting poor effort and feigning of cognitive impairment, but it may perform less well at detecting feigned psychopathology (Dandachi-FitzGerald & Merckelbach, 2013; Green et al., 2011). The MSVT is highly associated with the TOMM and has been found to produces similar rates of identifying suboptimal effort as the TOMM (Bashem et al., 2014).

Word Choice Test (WCT)

The Advanced Clinical Solutions (ACS) Word Choice Test (WCT) is a forced-choice performance validity test developed by Pearson in 2009. The WCT relies on recognition memory and consists of a 50-item word list. The administration consists of a learning trial and a test trial. During the learning trial, individuals are presented with the word list one word at a time. Each word is shown to the individual and is read out loud to them. In order to increase attention to each word, individuals are asked to indicate whether each word is something that is “man-made” or “natural”. After the learning trial, the test trial is immediately administered. The test trial consists of 50-word pairs, each containing a target word and a distractor. Individuals are asked to identify the target word in each pair. The WCT total score is based on the total correct items, with maximum score of 50. Administration takes approximately five to ten minutes. No universally accepted cut-off scores for the WCT have been established (Bain & Soble, 2019). For the purposes of the study, raw scores were utilized.

The initial validation sample consisted of individuals with diverse neurological and psychiatric disorders, including TBI, learning disorders, and ADHD, however it did not include those with probable dementia or Alzheimer’s Disease (Bain & Soble, 2019). Subsequent studies, however, have been conducted validating the WCT on various clinical samples, including those with neurocognitive disorders, epilepsy, Parkinson's Disease, frontotemportal lobar degeneration, and Alzheimer's disease (Bain & Soble, 2019). The WCT has been shown to have good specificity but weak sensitivity, and many studies have suggested that the WCT should not be used alone as a measure of performance validity (Bain & Soble, 2019; Armistead-Jehle & Buican, 2013).

Reliable Digit Span

The Reliable Digit Span (RDS) is a widely used embedded validity test. It is derived from the Digit Span subtest of the Weschler Intelligence Scales. It is calculated by summing the longest string of digits correctly repeated during the forward and backward trials of Digit Span, during which both trials were passed. RDS has been shown to have strong specificity and moderate sensitivity and has been shown to detect sub-optimal effort in those with TBIs, ADHD, psychotic disorders, and toxic exposure (Jasinski et al., 2011; Maiman et al., 2019). RDS has, however, been shown to have higher false positive rates in those with low IQ, severe memory disorders, strokes, and Non-English Native Speakers. Furthermore, ample evidence suggests that the Digit Span age-corrected scaled score may be a stronger PVT as compared to the RDS (Whitney et al., 2009). The typical cutoff score used for RDS is ≤ 6 or 7 (Maiman et al., 2019). For the purpose of this study, raw scores were utilized.

Minnesota Multiphasic Personality Inventory, Second Edition, Restructured Form

The Minnesota Multiphasic Personality Inventory, Second Edition, Restructured Form (MMPI-2-RF) is a new version of the MMPI-2 that was created by Ben-Porath and Tellegen in 2008. The MMPI-2-RF utilizes items from the MMPI-2; however, it is shorter, containing 338 items as compared to 567 items on the MMPI-2 (Van Der

Heijden et al., 2010). Additionally, the MMPI-2-RF utilizes the same normative sample as the MMPI-2, allowing for existing MMPI-2 data to be utilized as comparison groups and for research on the MMPI-2-RF (Van Der Heijden et al., 2010). Utilizing the same items and sample group as the MMPI-2 may serve as a disadvantage however, as items were last updated in the 1980's and may be outdated (Van Der Heijden et al., 2010).

The MMPI-2-RF may be administered in a paper/pencil or computerized format and takes approximately 35-50 minutes to complete. The MMPI-2-RF yields 51 scales, consisting of

nine validity scales, three higher order scales, nine restructured clinical scales, 23 specific problems scales, two interest scales, and five revised personality psychopathology (PSY–5) scales (Sellbom et al., 2018). Scales are arranged in a hierarchical fashion in terms of interpretation, with higher order scales being at the top and provide the most general indication of an individual's level of psychological difficulty (Forbey et al., 2010). Restructured scales are considered to be midlevel scales and specific problems and interest scales are considered at the bottom of the hierarchy and provide a narrow-band level of specific information (Forbey et al., 2010). These scales are supplemented by the personality psychopathology (PSY-5) scales, which represent a widely used and researched model of personality pathology (Sellbom et al., 2018). Raw scores and T-scores are provided for each scale. For the purpose of this study, T-scores will be utilized. Additionally, for the purpose of this study, the following MMPI-2-RF scales will be utilized as a measure of anxiety and depression: Demoralization (RCd), Low Positive Emotions (RC2), and Dysfunctional Negative Emotions (RC7). Evidence suggests that these scales are associated with internalizing problems and related psychopathology, specifically anxiety and depression (Wolf et al., 2008; Sellbom, 2017; Romero et al., 2017). As such, these scales were utilized as indicators of anxiety and depression for the current study.

Statistical Analysis

Prior to conducting the regression analyses, relevant regression assumptions were tested in order to assess skewness, kurtosis, multicollinearity, homoscedasticity, linearity, and normally distributed residuals of the sample and regression models. First, a descriptive analysis, including tests of skewness and kurtosis, was conducted. Any cases found to be influential over the models were removed from the dataset, so as to not distort the regression models and coefficient estimates. Additionally, homoscedasticity was assessed by plotting predicted values against standardized residual values and examining each scatterplot and histogram for an even distribution. Multicollinearity among the predictor variables were assessed by evaluating Pearson correlations and variance inflation factors (VIF).

To evaluate the hypothesis, hierarchical multiple regression analyses were conducted to evaluate the predicative ability of validity tests for memory performance. A regression analysis was conducted for each of the memory tests, for a total of six regression models. For each regression model, years played in the NFL and MMPI-2-RF Demoralization (RCd), Low Positive Emotions (RC2), and Dysfunctional Negative Emotions (RC7) were entered into the first block to assess their contribution to the models. Four validity measures (TOMM, MSVT, Word Choice, and RDS) were entered into the second block. Each memory subtest was entered as a dependent variable. Squared multiple correlation coefficients (R^2), as well as the change in multiple correlation coefficients (ΔR^2), were utilized to indicate the variance explained by the set of predictor variables. The p<.05 significance level was utilized for analyses.

While participants were administered three additional validity measures as part of their standardized neuropsychological evaluation (WMS Logical Memory, Visual Reproductions, and Verbal Paired Associates Recognition subtests), these subtests were excluded from the regression analyses given their association with the dependent variables. Given that the same stimulus items are utilized in the WMS subtests and their associated recognition tasks (i.e. same story is used in Logical Memory I, II, and Recognition; same word pairs used in Verbal Paired Associates I, II, and Recognition; same visual images used in Visual Reproductions I, II, and Recognition), and that all of the subtests are part of the same test (WMS-IV) and tap into similar memory constructs, the WMS-IV Recognition subtests were eliminated as predictor variables in order to reduce bias in the regression models. Eliminating WMS-IV Recognition subtests as predictors for all six regression analyses also allows for consistency across the regression models and better allows for direct comparison among the results. As a result, only four validity measures (TOMM, MSVT, Word Choice, and RDS) were utilized as predictors of the regression analyses.

Chapter IV: Results

Preliminary Results

Preliminary analyses were conducted to test statistical assumptions related to multiple regression models. Table 1 displays descriptive information about the variables used in the analyses. Based on Table 1, five variables were found to be leptokurtic (Logical Memory II, MSVT Immediate Recognition, MSVT Delayed Recognition, MMPI-2-RF RCd scale, and MMPI-2-RF RC2 scale). When examining the data for Logical Memory II and MMPI-2-RF RC2, one data point from each variable likely reflected a data entry error because it fell outside of possible values. As such, these two values were removed. The resulting descriptive information for Logical Memory II and MMPI-2-RF RC2 is displayed in Table 1 and is labeled LM II (adjusted) and MMPI-2-RF RC2 (Adjusted). After removing the outlier, skewness and kurtosis values for Logical Memory II and MMPI-2-RF RC2 were within the normal limits.

For regression analyses, non-normal distributions of variables are not uncommon and do not necessarily violate regression assumptions (Pituch & Stevens, 2016; Field, 2009). Rather, only the regression model residuals need to be normally distributed (Field, 2009). As such, further analyses were conducted to evaluate the residuals and determine if influential outliers exist for the remaining dataset. The homoscedasticity of the dataset, or the level of which residuals are equally distributed, was evaluated by examining P-P Plots and scatterplots for each of the six models. Scatterplots for each of the models appeared random and evenly distributed, indicating that the models did not violate the assumption of homoscedasticity. Additionally, all P-P Plots followed a linear fashion, illustrating the normality of the residuals. Histograms of the residuals were also used to assess for the normality of errors, which displayed normal distributions and did not indicate outliers.

Table 1

Descriptive Statistics for Neuropsychological Measures in the Sample

	N	Mean	SD	Skew	Kurtosis
LM I	126	6.67	3.18	.19	-.64
LM II	126	6.71	4.26	3.95	29.18
LM II (adjusted)	125	6.45	3.04	.44	.47
VPA I	126	7.32	2.63	.74	1.50
VPA II	126	7.50	2.83	.55	.30
VR I	126	7.92	3.30	-.14	-.59
VR II	126	9.03	2.97	.23	1.08
RDS	126	8.78	2.52	.52	1.28
Word Choice	126	43.60	7.38	-1.28	.62
TOMM Trial 2	126	44.98	7.84	-1.68	1.87
MSVT IR	126	95.64	10.00	-3.23	11.43
MSVT DR	126	90.83	13.50	-1.99	3.56
MSVT Consistency	126	90.39	13.03	-1.71	2.39
MSVT PA	126	79.76	23.41	-1.26	.87
MSVT FR	126	55.52	20.40	.01	-.55
Years in NFL	126	6.11	3.55	.90	1.52
RCd	126	52.77	3.73	-1.21	3.94
RC2	126	60.87	8.68	-2.01	11.87
RC2 (Adjusted)	125	61.31	7.17	-.10	.22
RC7	126	49.79	4.80	.17	1.19

Note: SD= Standard Deviation; RDS= Reliable Digit Span; LM= Logical Memory; VPA= Verbal Paired Associates; VR= Visual Reproductions; MSVT IR= MSVT Immediate Recognition; MSVT DR= MSVT Delayed Recognition; MSVT PA= MSVT

Paired Associates; MSVT FR= MSVT Free Recall; RCd= MMPI-2-RF Demoralization Restructured clinical scale; RC2= MMPI-2-RF Low Positive Emotions Restructured clinical scale; RC7= MMPI-2-RF Dysfunctional Negative Emotions Restructured clinical scale.

Multicollinearity among the predictor variables was assessed by evaluating Pearson correlations (Table 2) and the variance inflation factors (VIF; see regression analyses below). Variance inflation factors indicate whether there is substantial overlap among predictors (Pituch & Stevens, 2016). A VIF score is considered concerning if it is higher than 10 and variable deletion may be considered to modify the issue (Pituch & Stevens, 2016). Table 2 displays Pearson correlations for the predictor variables of the models. Table 2 indicates that MSVT Delayed Recognition and MSVT Consistency scores are

Table 2

Pearson Correlations between Predictor Variables

	1	2	3	4	5	6	7	8	9	10	11	12
1	-											
2	.43*	-										
3	.27*	.65*	-									
4	.35*	.66*	.78*	-								
5	.36*	.65*	.76*	.96*	-							
6	.45*	.54*	.56*	.64*	.67*	-						
7	.36*	.54*	.45*	.54*	.57*	.67*	-					
8	.37*	.74*	.61*	.75*	.73*	.54*	.44*	-				
9	-.18*	-.18*	-.22*	-.19*	-.18*	-.12	-.08	-.14	-			
10	.13	-.02	.05	-.01	-.03	-.01	-.11	.04	-.26*	-		
11	-.25*	-.18*	-.23*	-.24*	-.20*	-.23*	-.17	-.26*	.40*	-.41*	-	
12	.05	-.08	-.01	.00	-.02	-.09	-.16	-.07	-.08	.10	-.17	-

Note: 1=RDS; 2=WC; 3=MSVT IR; 4=MSVT DR; 5=MSVT Consistency; 6=MSVT PA; 7=MSVT FR; 8=TOMM; 9=MMPI-2-RF RCd; 10=MMPI-2-RF RC2; 11=MMPI-2-RF MC7; 12=Years played in NFL

* $p \leq .05$

highly correlated (r^2=.96, p≤.05). In order to diminish multicollinearity, MSVT Consistency was removed as a predictor. MSVT Delayed Recognition, as well Immediate Recognition, Paired Associates, and Free Recall are subtests of the MSVT and scores on these subtests are a direct result of an individual's performance. The MSVT Consistency variable is a computer-generated score derived from the MSVT scoring program (Suesse et al., 2015). Based on the Pearson correlations, MSVT Consistency is highly related to MSVT Delayed Recognition and as such, MSVT Consistency was removed as a predictor variable. After removing MSVT Consistency as a predictor, VIF values were examined for each of the regression models and were all under 10, suggesting that the multicollinearity assumption was met (see tables 3-8).

Overall, after removing two outlier data points and the MSVT Consistency variable as a predictor, these analyses indicated that the assumptions of multiple regressions, including multicollinearity, homoscedasticity, linearity, and normally distributed residuals were met across the regression models.

Results of Hypothesis

It was hypothesized that the four validity measures (TOMM, MSVT, RDS, and Word Choice) would significantly predict and account for performance on WMS-IV memory subtests (LM I and II, VPA I and II, VR I and II) when controlling for emotional functioning and years played in the NFL. In order to evaluate this hypothesis, six hierarchical multiple regression analyses were conducted to evaluate the predictive value of the validity measures above and beyond the effects of emotional functioning and the number of years players spent in the NFL.

Table three displays the multiple regression analyses results for predicting Logical Memory I performance. The model predicting Logical Memory I from the years in the NFL and the set of variables associated with emotional functioning alone was not significant (p=.361; R^2=.035). When the validity measure predictors were added to the regression model in a second step, the full model predicting Logical Memory I was statistically significant (p<.001). The validity measures accounted for about 40% of the variance in Logical Memory I, above and beyond the effects of emotional functioning and years played in the NFL, as indicated by the F-change test, ΔF (7, 113) = 9.85,

Table 3

Multiple Regression Analyses Predicting Logical Memory I

Predictor	B	SE	t	p	sr^2	VIF
Block 1	F (4, 120) = 1.10, p=.361, R^2=.035					
Years in NFL	0.00	0.07	0.04	0.969	0.00	1.09
MMPI RCd	-0.02	0.07	-0.35	0.729	0.00	1.26
MMPI RC2	-0.02	0.04	-0.54	0.590	0.00	1.30
MMPI RC7	0.06	0.06	0.96	0.340	0.00	1.53
Block 2	ΔF (7, 113) = 9.85, p<.001*, ΔR^2=.401					
RDS	0.23	0.11	2.05	0.042*	0.02	1.44
WC	-0.04	0.05	-0.79	0.430	0.00	2.98
MVST IR	0.01	0.04	0.31	0.754	0.00	2.89
MSVT DR	-0.01	0.04	-0.30	0.762	0.00	4.21
MSVT PA	0.01	0.02	0.56	0.578	0.00	2.46
MSVT FR	0.07	0.02	4.37	<.001*	0.10	2.12
TOMM	0.08	0.05	1.45	0.150	0.01	3.32

Note: Full model was statistically significant, F (13, 111) = 6.87, p<.001, R^2=.436

All coefficients are from final model.

*p<.05

p<.001, ΔR^2 =.401. Reliable Digit Span and MSVT Free Recall were significant predictors of Logical Memory I. Reliable Digit Span accounted for 2% of the variance of Logical Memory I and MSVT Free Recall accounted for 10% of the variance. The other validity measures were not found to be significant predictors of Logical Memory I (p >.05).

Table four displays the multiple regression analyses results for predicting Logical Memory II. The model predicting Logical Memory II from the years in the NFL and the set of variables associated with emotional functioning alone was not significant (p=.266; R^2 =.042). When the validity measure predictors were added to the regression model in a second step, the full model predicting Logical Memory II was statistically significant (p<.001). The validity measures accounted for about 36% of the variance in Logical Memory II, above and beyond the effects of emotional functioning and years played in the NFL, as indicated by the F-change test ΔF (7, 112) = 7.92, p<.001, ΔR^2 =.360. MSVT Free Recall was a significant predictor of Logical Memory II and accounted for 8% of the variance of Logical Memory II. The other validity measures were not found to be significant predictors of Logical Memory II (p >.05).

Table five displays the multiple regression analyses results for predicting Verbal Paired Associates I. The model predicting Verbal Paired Associates I from the years in the NFL and the set of variables associated with emotional functioning alone was not significant (p=.079; R^2 =.067). When the validity measure predictors were added to the regression model in a second step, the full model predicting Verbal Paired Associates I was statistically significant (p<.001). The validity measures accounted for about 35% of

Table 4

Multiple Regression Analyses Predicting Logical Memory II

Predictor	*B*	*SE*	*t*	*p*	sr^2	VIF
Block 1			*F* (4, 119) = 1.32, *p*=.266, R^2 =.042			
Years in NFL	-0.05	0.07	-0.69	0.490	0.00	1.10
MMPI RCd	-0.06	0.07	-0.82	0.412	0.00	1.26
MMPI RC2	0.00	0.04	0.01	0.996	0.00	1.30
MMPI RC7	0.05	0.06	0.88	0.381	0.00	1.52
Block 2			*ΔF* (7, 112) = 7.92, *p*<.001*, ΔR^2 =.360			
RDS	0.20	0.11	1.78	0.078	0.02	1.43
WC	-0.01	0.05	-0.16	0.876	0.00	2.97
MVST IR	-0.02	0.04	-0.49	0.624	0.00	2.89
MSVT DR	-0.01	0.04	-0.41	0.686	0.00	4.20
MSVT PA	0.01	0.02	0.91	0.367	0.00	2.47
MSVT FR	0.06	0.02	3.65	<.001*	0.08	2.12
TOMM	0.08	0.05	1.40	0.163	0.01	3.31

Note: Full model was statistically significant, *F* (11, 112) = 5.72, *p*<.001, R^2 =.402
All coefficients are from final model.
* p<.05

the variance in Verbal Paired Associates I, above and beyond the effects of emotional functioning and years played in the NFL, as indicated by the F-change test, *ΔF* (7, 113) = 7.08, *p*<.001, ΔR^2 =.351. Reliable Digit Span and MSVT Free Recall were significant predictors of Verbal Paired Associates I. Reliable Digit Span accounted for 5% of the variance of Verbal Paired Associates I and MSVT Free Recall accounted for 6% of the variance. The other validity measures were not found to be significant predictors of Verbal Paired Associates I (*p* >.05).

Table 5

Multiple Regression Analyses Predicting Verbal Paired Associates I

Predictor	B	SE	t	p	sr^2	VIF
Block 1			F (4, 120) = 2.15, p=.079, R^2=.067			
Years in NFL	-0.07	0.06	-1.17	0.247	0.01	1.09
MMPI RCd	-0.03	0.06	-0.51	0.614	0.00	1.26
MMPI RC2	-0.01	0.03	-0.36	0.720	0.00	1.30
MMPI RC7	-0.02	0.05	-0.44	0.663	0.00	1.53
Block 2			ΔF (7, 113) = 7.08, p<.001*, ΔR^2=.351			
RDS	0.27	0.10	2.84	0.005*	0.05	1.44
WC	0.08	0.05	1.75	0.082	0.02	2.98
MVST IR	0.02	0.03	0.52	0.601	0.00	2.89
MSVT DR	-0.02	0.03	-0.53	0.599	0.00	4.21
MSVT PA	-0.01	0.01	-0.99	0.324	0.01	2.46
MSVT FR	0.05	0.01	3.21	0.002*	0.06	2.12
TOMM	-0.01	0.05	-0.28	0.780	0.00	3.32

Note: Full model was statistically significant, F (11, 113) = 5.57, p<.001, R^2=.418
All coefficients are from final model.
*p<.05

Table six displays the multiple regression analyses results for predicting Verbal Paired Associates II. The overall model predicting Verbal Paired Associates II from the years in the NFL and the set of variables associated with emotional functioning alone was not significant (p=.057; R^2=.073). When the validity measure predictors were added to the regression model in a second step, the full model predicting Verbal Paired Associates II was statistically significant (p<.001). The validity measures accounted for about 41%

Table 6

Multiple Regression Analyses Predicting Verbal Paired Associates II

Predictor	*B*	*SE*	*t*	*p*	sr^2	VIF
Block 1	F (4, 120) = 2.36, p=.057, R^2=.073					
Years in NFL	-0.10	0.06	-1.66	0.101	0.01	1.09
MMPI RCd	0.02	0.06	0.25	0.806	0.00	1.26
MMPI RC2	-0.03	0.03	-0.78	0.436	0.00	1.30
MMPI RC7	0.02	0.05	0.39	0.700	0.00	1.53
Block 2	ΔF (7, 113) = 9.25, p<.001*, ΔR^2=.411					
RDS	0.15	0.10	1.55	0.124	0.01	1.44
WC	0.11	0.05	2.33	0.022*	0.03	2.98
MVST IR	-0.04	0.04	-1.01	0.313	0.01	2.89
MSVT DR	0.00	0.03	0.08	0.938	0.00	4.21
MSVT PA	-0.02	0.01	-1.54	0.126	0.01	2.46
MSVT FR	0.06	0.02	4.16	<.001*	0.09	2.12
TOMM	0.04	0.05	0.79	0.429	0.00	3.32

Note: Full model was statistically significant, F (11, 113) = 7.16, p<.001, R^2=.484
All coefficients are from final model.
* p<.05

of the variance in Verbal Paired Associates II, above and beyond the effects of emotional functioning and years played in the NFL, as indicated by the F-change test ΔF (7, 113) = 9.25, p<.001, ΔR^2=.411. MSVT Free Recall and Word Choice were significant predictors of Verbal Paired Associates II. MSVT Free Recall accounted for 9% of the variance of Verbal Paired Associates II and Word Choice accounted for 3%. The other validity measures were not found to be significant predictors of Verbal Paired Associates II (p>.05).

Table seven displays the multiple regression analyses results for predicting Visual Reproductions I. The overall model predicting Visual Reproductions I from the years in the NFL and the set of variables associated with emotional functioning alone was significant (p=.044). This model accounted for about 8% of the variance in Visual Reproductions I, F (4, 120) = 2.53, p=.044, R^2=.078. After adding the validity measures into the model, the full model predicting Visual Reproductions I was statistically significant (p<.001). The validity measures accounted for an additional 40% of the variance in Visual Reproductions I, above and beyond the effects of emotional

Table 7

Multiple Regression Analyses Predicting Visual Reproductions I

Predictor	*B*	*SE*	*t*	*p*	sr^2	VIF
Block 1	F (4, 120) = 2.53, p=.044*, R^2=.078					
Years in NFL	0.02	0.07	0.21	0.833	0.00	1.09
MMPI RCd	-0.05	0.07	-0.68	0.501	0.00	1.26
MMPI RC2	-0.07	0.04	-1.92	0.057	0.02	1.30
MMPI RC7	-0.01	0.06	-0.22	0.825	0.00	1.53
Block 2	ΔF (7, 113) = 8.49, p<.001*, ΔR^2=.396					
RDS	0.26	0.12	2.28	0.024*	0.03	1.44
WC	-0.04	0.06	-0.75	0.458	0.00	2.98
MVST IR	0.01	0.04	0.12	0.906	0.00	2.89
MSVT DR	0.02	0.04	0.59	0.557	0.00	4.21
MSVT PA	-0.02	0.02	-1.37	0.173	0.01	2.46
MSVT FR	0.06	0.02	3.39	0.001*	0.06	2.12
TOMM	0.14	0.06	2.49	0.014*	0.03	3.32

Note: Full model was statistically significant, F (11, 113) = 6.72, p<.001, R^2=.474

All coefficients are from final model.

* p<.05

functioning and years played in the NFL, as indicated by the F-change test, ΔF (7, 113) = 8.49, $p<.001$, $\Delta R^2=.396$. Reliable Digit Span, MSVT Free Recall, and TOMM were significant predictors of Visual Reproductions I. Reliable Digit Span accounted for 3% of the variance, MSVT Free Recall accounted for 6% of the variance, and TOMM accounted for 3% of the variance. The other validity measures were not found to be significant predictors of Visual Reproductions I ($p>.05$).

Table eight displays the multiple regression analyses results for predicting Visual Reproductions II. The overall model predicting Visual Reproductions II from the years in the NFL and the set of variables associated with emotional functioning alone was significant ($p=.018$). This model accounted for about 9% of the variance in Visual Reproductions II, F (4, 120) = 3.10, $p=.018$, $R^2=.094$. After adding the validity measures into the model, the full model predicting Visual Reproductions II was statistically significant ($p<.001$). The validity measures accounted for an additional 49% of the variance in Visual Reproductions II, above and beyond the effects of emotional functioning and years played in the NFL, as indicated by the F-change test, ΔF (7, 113) = 12.71 $p<.001$, $\Delta R^2=.493$. MMPI-2-RF RC2, MSVT Free Recall, and TOMM were significant predictors of Visual Reproductions II. MMPI-2-RF RC2 accounted for 2% of the variance of Visual Reproductions II, MSVT Free Recall accounted for 10% of the variance, and TOMM accounted for 2% of the variance. The other validity measures were not found to be significant predictors of Visual Reproductions II (p >.05).

Table 8

Multiple Regression Analyses Predicting Visual Reproductions II

Predictor	B	SE	t	p	sr^2	VIF
Block 1	F (4, 120) = 3.10, p=.018*, R^2=.094					
Years in NFL	0.04	0.06	0.73	0.467	0.00	1.09
MMPI RCd	0.07	0.06	1.12	0.264	0.01	1.26
MMPI RC2	-0.07	0.03	-2.35	0.020*	0.02	1.30
MMPI RC7	-0.05	0.05	-0.90	0.368	0.00	1.53
Block 2	ΔF (7, 113) = 12.71, p<.001*, ΔR^2=.493					
RDS	0.14	0.09	1.45	0.149	0.01	1.44
WC	0.07	0.05	1.41	0.161	0.01	2.98
MVST IR	0.01	0.03	0.42	0.679	0.00	2.89
MSVT DR	-0.03	0.03	-1.09	0.279	0.01	4.21
MSVT PA	-0.01	0.01	-0.89	0.375	0.00	2.46
MSVT FR	0.07	0.01	4.73	<.001*	0.10	2.12
TOMM	0.10	0.05	2.19	0.031*	0.02	3.32

Note: Full model was statistically significant, F (11, 113) = 9.99, p<.001, R^2=.587
All coefficients are from final model.
*p<.05

The present results support the hypothesis. Each of the six regression analyses yielded significant regression models and demonstrated that the validity measures accounted for a significant amount of the variance of memory performance above and beyond the effects of emotional functioning and years played in the NFL. These findings demonstrate the unique variance in memory performance accounted for by validity measures.

Chapter V: Discussion

Discussion of Results

The purpose of this study was to examine the relationship between validity measures and memory measures in retired NFL players and to explore how predictive validity measures were for memory performance in this population. Research regarding long-term outcomes of repeated concussions from football has been well established, however no research has examined the validity of this currently used battery in retired players. Currently, some research exists examining effort and validity testing in athletes, however much of the focus of this literature is on the underreporting of symptom rather than on neuropsychological performance. Additionally, much of the research focuses on high school and college level athletes, and few studies examine professional athletes, such as NFL players. As such, this study attempted to provide relevant clinical information regarding the assessment of a unique population, retired NFL players. The purpose of examining the relationship between the validity measures and memory measures in this battery was to evaluate the extent to which validity measures are associated with and/or predict memory performance. This study may provide important information regarding decisions made by the NFL concussion settlement, whether these validity measures provide valid information regarding neuropsychological performance, and what it means to meet or not meet criteria for cognitive impairment.

It was hypothesized that four validity measures (TOMM, MSVT, RDS, and Word Choice) would significantly predict and account for performance on WMS-IV memory subtests (Logical Memory I and II, Visual Reproductions I and II, Verbal Paired Associates I and II) when controlling for emotional functioning and years played in the

NFL. In order to evaluate this hypothesis, hierarchical multiple regression analyses were conducted to evaluate the predicative ability of validity tests for memory performance. A regression analysis was conducted for each of the memory tests, for a total of six regression models.

The hypothesis that the four validity measures would significantly predict and account for performance on WMS-IV memory subtests was supported by the present results. Each of the six regression analyses yielded significant regression models and demonstrated that the validity measures accounted for a significant amount of the variance of memory performance above and beyond the effects of emotional functioning and years played in the NFL. These findings demonstrate the unique variance in memory performance accounted for by validity measures.

Based on the current results, MSVT Free Recall (FR) appeared to be the most robust predictor of memory performance in this sample, as MSVT Free Recall was found to be a significant predictor across the six regression models. The other MSVT variables (Immediate Recognition (IR), Delayed Recognition (DR), and Paired Associates (PA)) were found to be poor predictors of memory performance, as none of the variables were significant predictors of memory performance across the regression models. The unique contribution accounted for by the MSVT Free Recall subtest may be related to the underlying memory component associated with the subtest. IR and DR are considered to be the primary effort subtests of the MSVT, while PA and FR are considered to be measures of memory, with FR being the most difficult subtest of the MSVT (Green, 2004). Furthermore, when interpreting the FR subtest, the subtest should only be interpreted after examining the validity of the IR, DR, Consistency, and PA subtests, and then the FR score may be interpreted similarly to that of any verbal recall test (Green, 2004). This further suggests that the FR score is not reflective of a measure of effort or validity and rather functions as a measure of memory.

Armstead-Jehle and Hansen (2016) conducted a factor analysis suggesting that the MSVT FR serves more as a measure of memory than an effort measure. They conducted a factor analysis of the MSVT, the Nonverbal MSVT (NV-MSVT), and the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS) memory subtests, and found that the analyses yielded two factors, one which was labeled effort and one which was labeled memory. All RBANS memory subtests, as well as the FR subtest from both the MSVT and NV-MSVT were loaded onto the memory factor and the remaining MSVT and NV-MSVT variables were loaded onto the effort factor, suggesting that the FR subtest functions more as a measure of true memory ability rather than as a measure of effort or validity (Armstead-Jehle & Hansen, 2016). This further suggests that the memory construct incorporated into the MSVT FR subtest may be accounting for the significant relationship found between the subtest and the memory outcome variables in the present study.

Given that many of the NFL players in the current study presented with severe cognitive complaints, it was expected that many of the participants would demonstrate poor scores on the FR subtest (M=55.52, SD=20.40) and the PA subtest (M=79.76, SD=23.41). However poor scores on the FR or PA subtests do not necessarily imply poor effort and could be a result of true memory impairment (Green, 2004). Furthermore, individuals with severe cognitive impairment may also be incapable of passing the effort subtests of the MSVT using the typical cut-off score of 85%. As such, Green (2004) created a separate set of criteria to evaluate valid MSVT scores in those with severe cognitive impairment and who likely have a profile consist with dementia. The criteria consists of the following: failure on IR, DR or Consistency Scores; the mean of IR, DR, and Consistency scores should be at least 20 points higher than the mean of PA and FR scores; PA score should be higher than FR score; and the IR and DR score should be higher than FR. This criterion may provide more clinically meaningful information and reduce false positive rates of poor effort in retired NFL players presenting with severe cognitive impairment.

Reliable Digit Span (RDS) was found to be a significant predictor of the immediate memory subtests (Logical Memory I, Verbal Paired Associates I, and Visual Reproductions I) and was non-significant for delayed memory subtests (Logical Memory II, Verbal Paired Associates II, and Visual Reproductions II). RDS is one of the most widely used embedded validity measures and has been found to detect suboptimal effort in several clinical populations (Maiman et al., 2019). Its use as a performance validity test has been questionable however, particularly when being used with individuals with memory disorders/dementias and when utilizing the typical cut-off score of less than 6 or 7 (Maiman et al., 2019; Kiewel et al., 2012; Schroeder et al., 2012). Even when using a cut-off of 5, false positive rates of invalid test performance has been found to be 18% among memory impaired individuals (Loring et al., 2016). Furthermore, the normative data provided by the Advanced Clinical Solutions (ACS) manual, which provides administration and scoring information on commonly used embedded Performance Validity Tests (PVTs), including RDS, does not include normative data on those with dementia. This limits the generalizability of the scoring system in those with severe cognitive impairment, similar to the NFL sample utilized in this study. Rather, the ACS normative data includes the following clinical groups: TBI, temporal lobectomy, schizophrenia, major depressive disorder, anxiety, mild intellectual disability, autistic disorder, Asperger's Disorder, Reading Disorder, Math Disorder, and ADHD (Wechsler, 2009).

RDS is derived from the Digit Span subtest of the Weschler Scales, which is a subtest measuring attention, auditory processing, and mental manipulation. As such, RDS has been found to be significantly associated with intelligence and attention, as well as a list learning task (Maiman et al., 2019). On the same note, RDS has also been found to have a non-significant relationship with well-known performance validity measures, such as the TOMM, which further suggests that the RDS may function more as a measure of attention rather than as a performance validity test (Maiman et al., 2019).

In the current results, RDS was observed to be a significant predictor of Logical Memory I, Verbal Paired Associates I, and Visual Reproductions I. RDS was not found to be a significant predictor for any of the three delayed memory tests. Attention and working memory have been implicated in memory, particularly in terms of memory encoding (Chun & Turk-Brown, 2007). This is because attentional processes help us selectively attend to information to be perceived, encoded, and stored for later retrieval. Working memory is used to carry out and plan behavior, and is related to processing, manipulation, and attentional control (Cowan, 2008). Immediate memory tasks, such as those on the WMS-IV, utilize what is called short-term memory. While some researchers distinguish working memory and short-term memory, the processes are not completely distinct, as they both reflect processes that can hold a limited amount of information for a short period of time (Cowan, 2008). Differences in definitions may exist based on the theoretical model of memory being used or what processes are being studied, such as manipulation, attentional control, or the relationship with aptitude or intelligence (Cowan, 2008). Given the relationship between working memory, encoding, and short-term memory, it is expected that RDS would be associated with immediate memory tasks, as these tasks require the utilization of working memory processes to attend to, learn, and encode the information presented in these subtests.

Delayed memory tasks, such as those on the WMS-IV, involve long-term memory. While different theoretical models of memory exist, most models differentiate long-term memory as a distinct category of memory. Historically, William James (1890) referred to immediate memory as primary memory and other memory (including long-term) as secondary memory (Kellogg, 1997). Waugh and Norman (1965) continued based off James' work and formalized a model of separate primary and secondary memories. In Atkinson and Shiffrin's Multistore Model (1971), memory is broken down into sensory, short-term, and long-term memory. In Baddeley and Hitch's Working Memory Model (1974), memory is divided into sensory memory, central executive, visuo-spatial scratch pad, phonological loop, and long-term memory.

Short-term and long-term memory can be differentiated most notably by capacity (how many items can be stored) and duration (how long items can be stored), however retrieval process of short-term and long-term memory also differ, with short-term memory using a serial exhaustive approach (items in memory are all examined one by one) and long-term memory utilizing a parallel approach, such that all items are examined simultaneously (Kellogg, 1997; Cowan, 2008). Evidence of distinctions between short-term and long-term memory processes have been demonstrated in neuropsychological research. For example, in a study examining how Auditory Verbal Learning Test (AVLT) scores can be predicted from MRI results, Moradi and colleagues (2017) found that the top predictors of immediate memory scores were right middle temporal gyrus, right amygdala, left insula and the top predictors of delayed memory were the

right angular gyrus and the bilateral hippocampus. Additionally, in an fMRI study on memory processes for digits, Nie et al. (2019) found that short-term memory for the digits were located in the visual cortex and were encoded by visual representations and that long-term memory was encoded by semantics and utilized the left frontal cortex, suggesting that different forms of memory utilize different neural and mental processes.

As such, it is possible that in the current results, RDS was not found to be associated with delayed memory tasks given the distinct neurological processes involved in short-term and long-term memory. The underlying working memory construct incorporated into RDS may be accounting for the significant relationship found between RDS and the immediate memory outcome variables in the present study. Furthermore, if the participants were truly forgetting information that they learned or were displaying deficits in the ability to independently recall information, then RDS may not be related to delayed memory subtests, as this would suggest impairment in memory retrieval rather than memory encoding.

In terms of the TOMM, the current results demonstrated that the TOMM was significantly associated with Visual Reproductions I and II. Similar to the MSVT Free Recall and RDS results, the unique contribution accounted for by the TOMM may be related to the underlying visual recognition component associated with the test. The TOMM relies on visual attention and recognition to complete the task. Similarly, Visual Reproductions I and II rely on visual attention and memory processes. While the TOMM was designed to be a relatively easy test even for those with intellectual or neurological impairment, high false positive rates have been demonstrated in those with severe cognitive impairment, particularly when using the typical cutoff score of 45 or greater (Tombaugh, 1997; Dean et al., 2009).

The original normative data of the TOMM included the following groups: no cognitive impairment, cognitive impairment (including amnesia, multiple sclerosis, Parkinson’s disease, substance use, stroke, ADHD, learning disability, depression, Huntington’s disease, subdural hematoma, carbon monoxide exposure, arteriovenous malformation rupture, colloid cyst, and other), TBI, and dementia (Tombaugh, 1997). While a TBI group was included in the development of the TOMM, the severity of the TBIs were not noted, making it difficult to discern the generalizability of the test and cut-off score of 45 to those with moderate/severe or repeated head injuries, such as those in our current sample. Furthermore, when examining those with dementia, 27% of dementia patients scored below 45 (Tombaugh, 1997). This further suggests that the standard cut-off may not be suitable for those with severe cognitive impairment. Moreover, the majority of the sample in the TOMM normative data consisted of neurologically intact individuals, as three out of the four original experiments were conducted with only neurologically intact individuals and only one study utilized neurologically-impaired individuals.

While various forms of memory have been found to be impacted in those with TBIs, including verbal and visual memory, learning, and prospective memory, visual memory has been found to be more impacted than auditory memory (Vakil, 2005; Carlozzi et al., 2013). For example, individuals with complicated mild/moderate TBI have been found to perform poorly on WMS-IV subtests of visual memory and visual working memory as compared to controls, but not on auditory memory tasks (Carlozzi, et al., 2013). As such, increased impairment on visual attention and memory may contribute to difficulties on the TOMM, such that if an individual is not able to attend to the stimulus presented during administration, he/she will not be able to complete the task adequately. Given that many of the NFL players in the sample presented with attention and memory complaints secondary to their head injuries, it is likely that visual attention/memory difficulties may have contributed to difficulties on this task. Furthermore, this may have also contributed to the significant relationship between TOMM and Visual Reproductions performance as indicated in the current results, further suggesting that the relationship may be due to underlying visual attention deficit rather than poor effort. As such, the TOMM may not be adequate for those with severe visual attention or memory impairment, such as the NFL players in the current sample. Rather, lower cutoff scores on the TOMM may be necessary to avoid false positive rates of poor effort. Additionally, the use of multiple PVTs including those that do not rely on visual recognition/attention, may be useful in order to get accurate representation of true cognitive performance.

Another pattern displayed in the current results was that block one was significant for only Visual Reproductions I and II. In block one, MMPI-2-RF RC2 and RC7 were clinically significant predictors for both Visual Reproductions I and II, however after adding the remaining predictors in block two, only RC2 was significant for Visual Reproductions II. The RC2 scale of the MMPI-2-RF is the Low Positive Emotions restructured scale and is associated with a lack of positive emotional experiences, anhedonia, and most notably depressive disorders, however this scale may also be elevated in those with schizophrenia or PTSD (Ben-Porath, 2012; Sellbom, 2017). The RC7 is the Dysfunctional Negative Emotions restructured scale and is associated with anxiety, anger, fear, and irritability, such that high scores on this scale is associated with anxiety disorders (Ben-Porath, 2012). As previously mentioned, evidence exists to suggest that visual memory may be more impacted than auditory memory in those with TBIs (Vakil, 2005; Carlozzi et al., 2013). As such, it is possible that visual memory tasks may be more challenging for this population and may elicit a stronger emotional response when completing these tasks. This could also be interpreted such that increased negative emotionality may interfere more with visual memory tasks, given that this population may already have a disadvantage with visual memory performance.

In the current results, some well-known PVTs were found to be poor predictors of memory, including the TOMM, MSVT effort subtests, and ACS Word Choice. PVTs including the TOMM and MSVT, were designed to detect feigning, malingering, and exaggeration of cognitive and other symptoms (Dandachi-FitzGerald & Merckelbach, 2013; Green, 2004; Tombaugh 1997). They function such that failure on PVTs is suggestive of potential feigning, and many validation studies of PVTs demonstrate that poor performance on PVTs is predictive of poor performance on cognitive tests. For example, in a study on neuropsychological performance and effort in Gulf War era veterans, Lindem et al. (2003) found that low scores on the TOMM were associated with low scores on tests of attention, executive functioning, and memory. Similar results have also been found on self-report measures of cognitive symptoms. For example, individuals who failed the MSVT were found to display significantly more memory complaints than those who passed the MSVT, suggesting that those who display poor effort also exaggerate memory symptoms on self-report measures (Green, 2004).

While poor performance on PVTs have been associated with poor cognitive performance, the opposite is not necessarily true- passing PVTs does not necessarily indicate strong performance on cognitive tests. Given that the purpose of PVTs is to differentiate genuine versus feigned symptoms, it is possible for individuals to pass PVTs but still perform poorly on cognitive tests. This profile is suggestive of good effort and true cognitive deficits. As such, the current results may not have yielded a significant relationship between PVTs and memory performance if participants passed PVTs and also did well on cognitive tests, or if participants passed PVTs but did poorly on memory tests.

Overall, the current results demonstrated significant relationships between PVTs and memory performance in the present clinical battery, however the nature of these relationships may be impacted by underlying cognitive constructs, rather than motivation or true effort put forth on performance. These results provide clinically meaningful information and should be considered when interpreting PVTs as part of this clinical battery for retired NFL players, as many of these individuals present with severe cognitive impairment.

Theoretical and Clinical Implications

The current study encompasses several implications, both in terms of theory and clinical practice. One of the most salient clinical applications is that standard cut-off scores utilized for various PVTs may not be valid or applicable to all clinical populations. This is particularly true for those who display severe cognitive impairment, TBIs, and/or dementia, as it has been demonstrated that failure on PVTs does not necessarily suggest poor effort and could be a result of true cognitive impairment (Green, 2004; McWhirter et al., 2020). This suggests that when evaluating cognitive impairment and performance on PVTs, clinicians must consider whether or not to use traditional cut-off scores. By utilizing lower cut-off scores on PVTs, clinicians can help to reduce false positive PVT scores in those who are vulnerable to false positive profiles. As such, this study adds to the existing literature that PVTs, as traditionally used, may not be valid for those with severe cognitive impairment and that utilizing traditional cut-off scores may lead to the misclassification of malingering.

Substantial evidence exists to suggest that lowered cut-off scores may be appropriate, particularly when evaluating those with TBIs, dementia, MCI, or other forms of severe cognitive impairment (McWhirter, et al., 2020; Dean et al., 2009; Teichner and Wagner, 2004). This concept has been suggested for various PVTs, including the RDS, MSVT, and the TOMM. This has also been suggested for forced choice tests generally, which is a common paradigm utilized for PVTs, as utilizing higher than chance cut-offs may make the test redundant and susceptible to those with any sort of attentional deficit (McWhirter, et al., 2020). Dean and colleagues (2009) conducted a study in which they evaluated the validity of various stand-alone and embedded PVTs in a sample of dementia participants, using published cut-off scores. They found that several commonly used PVTs including the TOMM, RDS, Rey 15, and other PVTs had specificity rates of 70% or lower, with the lowest being 13% (RAVLT effort equation). They concluded that commonly used PVTs are unacceptable for assessing those with dementia when using traditional cut-off scores and that lowered scores should be utilized when assessing this population (Dean et al., 2009). In a study evaluating the use of the TOMM in those with dementia, Teichner and Wagner (2004) found that those who had dementia were misclassified as malingering when using the traditional cut-off score of five errors, as well as when using cut-off scores of eight or ten. They concluded that the TOMM can be useful for assessing malingering, however only after dementia is ruled out (Teichner and Wagner, 2004). Overall, the current study supports the idea that lowered cut-off scores for PVTs should be considered for those with severe cognitive impairment, so as to reduce the misclassification of malingering in these populations.

Another important implication of the current study is that caution should be taken when utilizing validity tests that incorporate underlying cognitive constructs. While failure on PVTs is typically thought to indicate suboptimal effort, cognitive impairment can make it difficult to successfully complete PVTs. This is particularly true for embedded validity tests, such as the Reliable Digit Span, which are sensitive to genuine cognitive impairment (Zenisek et al., 2016). This is also true for tests that rely on memory processes, such as the MSVT Free Recall subtest (Green, 2004). This concept was supported by the current results, as the current results suggested that the predictive ability of the RDS, MSVT Free Recall, and the TOMM for memory performance was influenced by attention and memory processes. Given that these and other PVTs may be sensitive to cognitive functioning, poor performance on these tasks can easily be mistaken for suboptimal effort or malingering. As such, it is important to evaluate the validity and specific characteristics of PVTs and their use for evaluating those with dementia and cognitive impairment, as higher false-positive rates associated with higher cognitive impairment may suggest that these tests are inappropriate to use when evaluating these populations (Zenisek et al., 2016).

PVTs exist across multiple domains, including intelligence, memory, processing speed, attention, visual-spatial, language, executive functioning, and motor skills (Sweet et al., 2021). One way of reducing bias related to cognitive functioning and PVTs is by utilizing multiple PVTs that cover multiple domains. The American Academy of Clinical Neuropsychology (AACN) recently published a consensus statement on validity assessment. The statement emphasized the use of multiple PVTs, as well as the use of both stand-alone and embedded PVTs. They also argue for the use of PVTs that cover multiple cognitive domains. By utilizing multiple PVTs that incorporate different cognitive domains, clinicians may improve the validity of neuropsychological data interpretation, reduce bias that may exist from true cognitive impairment, and assess for potential feigning across specific domains.

An important topic explored in the current study was the impact of mood and emotional distress on cognitive performance. The current results suggest that mood difficulties may exacerbate poor cognitive performance, particularly on tasks that are more difficult for individuals. The effects of emotional distress on neuropsychological performance may exist for a variety of reasons. Effects of emotional distress may be due to related mood symptoms, such as psychomotor retardation, fatigue, reduced motivation, and general cognitive inefficiency, all of which have been implicated in anxiety and depression (Gass, 1991; Lezak et al., 2012). Additionally, negative expectations about one's capabilities may impact performance. For example, high levels of anxiety and depression have been associated with slowed thinking and responding, mental blocks, distractibility and difficulties with memory (Lezak et al., 2012). It is also important to consider, however, that anxiety and depression are common consequences of TBIs. This common occurrence has several important implications.

One important implication is that psychiatric comorbidities may complicate cognitive evaluations. Symptoms of the psychiatric comorbidities may worsen cognitive performance, as mentioned above. Psychiatric comorbidities can also make it difficult to differentiate cognitive, somatic, and behavioral symptoms from those that are directly caused by the TBI (Schwarzbold, et al., 2008). Furthermore, it is possible that comorbid TBI and psychiatric disorders can make assessing effort and performance validity challenging, and that the effects of emotional distress may be mistaken for poor effort. Emotional distress may impact motivation and engagement in testing, which in turn may impact performance validity. Additionally, given that many psychiatric disorders, such as anxiety and depression, are associated with attention difficulties and other cognitive symptoms, it is possible that comorbid psychiatric disorders may contribute to high PVT failure rates (Bigler, 2014). Take for example, an individual who is presenting with attentional difficulties secondary to a TBI and is also depressed and having difficulties with motivation and allocating cognitive resources to engage in tasks. The patient would be more likely to fail tests like the TOMM or RDS that rely on one's ability to actively engage in the task and attend to stimuli being presented. In other words, if an individual is not paying attention during PVTs because he/she is unmotivated or distracted by emotional symptoms, negative cognitive beliefs, or other symptoms related to the comorbid psychiatric disorder, then the individual may not be able to successfully complete the task and subsequently exhibit a false positive PVT result. As such, given that those with TBIs commonly fail PVTs and frequently demonstrate comorbid anxiety and/or depression, the effects of these psychiatric disorders need to be considered as a potential contributing factor to poor performance, in addition to cognitive deficits and possible intentional feigning.

The effects of psychiatric conditions on cognitive performance have been demonstrated in the literature. For example, Rapoport and colleagues (2005) conducted a study in which they explored the relationship between depression and cognitive impairment following TBI. They found that TBI patients with comorbid depression performed worse on tasks of working memory, processing speed, verbal memory, and executive functioning, as compared to TBI patients without depression. Similarly, Mohn and Rund (2016) conducted a study on the impact of depression on cognitive performance and found those with depression to perform significantly worse on measures of working memory, processing speed, attention, learning, and problem solving as compared to those without depression. Additionally, addressing issues related to motivation has been found to improve performance. For instance, it has been found that confronting non-litigating individuals who initially display poor performance on PVTs results in higher scores after re-assessment, suggesting that they were initially not fully engaged in performance (Bigler, 2014).

In terms of anxiety/depression and PVTs, many studies examining the effects of anxiety and depression on PVTs demonstrate that these psychiatric disorders alone do not commonly result in false positive results (Boone, 2021). Other psychiatric disorders have been shown to impact performance validity, including psychotic disorders and PTSD (Boone, 2021; Jurick, et al., 2019). Little research, however, has been conducted on the effects of comorbid TBI and psychiatric disorders on PVTs. In one study on the effects of psychological treatment for those with comorbid TBI and PTSD, it was found that psychological treatment improved PVT performance and reduced symptoms, suggesting a relationship between psychological distress and PVT performance (Jurick, et al., 2019). The current results suggest that emotional distress may exacerbate poor performance on cognitive tasks or PVTs and that the effects of comorbid anxiety/depression is an important issue that needs to be considered when evaluating those with cognitive impairment. This adds to the gap in the literature in terms of exploring the relationship between cognitive impairment, emotional distress, and performance on PVTs.

In regard to assessing retired NFL players, the current results suggest some PVTs utilized in the current battery may not be valid to use for evaluating this population. This is particularly true for players displaying high levels of cognitive impairment and/or comorbid anxiety/depression, when traditional cut-off scores for PVTs are utilized, or when utilizing PVTs that rely on cognitive functioning. The results imply that bias may exist, as much of the samples used to validate PVTs often exclude those with dementia/severe cognitive impairment (Zenisek et al., 2016). This serves as an important clinical consideration, as many of the retired NFL players presented with, and were being evaluated for, severe cognitive impairment. As such, caution must be taken when evaluating cognitive performance and performance on PVTs in order to differentiate between genuine cognitive impairment, emotional distress, and suboptimal effort.

Several steps and considerations may be taken to improve the validity of these evaluations. This may include utilizing lower cut-off scores for PVTs and multiple PVTs that encompass several cognitive domains, as outlined previously. This may also include considering

the various potential causes of PVT failure. As mentioned previously, genuine cognitive impairment and emotional distress may contribute to PVT failure, however other causes of PVT failure may include intentional feigning, apathy, lack of motivation, fatigue, or pain, all of which should be considered when interpreting PVT performance (McWhirter et al., 2020). Additionally, multiple sources of information should be utilized to assess test validity, which may include reviewing medical records, academic records, and previous neuropsychological assessments; conducting a comprehensive clinical interview; and utilizing Symptom Validity Tests (SVTs) in addition to PVTs (Sweet et al., 2021). Moreover, clinical judgement should be utilized when interpreting invalidity and clinicians should not rely solely on whether or not an individual passes PVTs (Sweet et al., 2021). All in all, caution must be taken when interpreting neuropsychological results of this unique population and what it means in terms of their cognitive functioning and overall performance.

Overall, the current study provided important clinical and theoretical implications. Namely, that standard cut-off scores utilized for various PVTs may not be valid or applicable to all clinical populations, particularly those displaying TBIs, dementia, or other forms of severe cognitive impairment. This is especially true for tests that rely on cognitive functioning or that are embedded as part of cognitive tests, as they may be sensitive to genuine cognitive impairment. Emotional difficulties and comorbid psychiatric disorders are important factors to consider when interpreting cognitive performance and PVTs, as mood difficulties may exacerbate poor performance on these tests. As a result, caution must be taken when evaluating cognitive performance and performance on PVTs in order to differentiate between genuine cognitive impairment, emotional distress, and suboptimal effort. This may include considering different factors that may contribute to failure on PVTs, carefully considering what cut-off scores to use for PVTs, utilizing multiple PVTs that cover different cognitive domains, using multiple sources of information, and utilizing clinical judgement when interpreting invalidity and not relying solely on whether or not an individual passes PVTs.

In addition, this study provided important clinical implications for evaluating retired NFL players. This study provided more information regarding the relationship between validity measures and memory performance in the current battery. More specifically, it helped to elucidate whether these validity measures provide valid information regarding neuropsychological performance for these players and what it means to meet or not meet criteria for cognitive impairment. It also addressed the lapse in information regarding the use of these various validity measures in providing neuropsychological assessments for retired NFL players. These findings may be used to provide important information regarding decisions made by the NFL concussion settlement and can help clinicians make better decisions regarding client functioning and treatment.

Integration with Literature

The findings and implications of the current study contribute to the growing literature regarding validity testing with neurologically impaired individuals. Substantial controversy surrounds the use of PVTs for evaluating those with TBIs, as generally PVTs are designed to be relatively easy tests, even for those with intellectual or neurological impairment. Moreover, most individuals with a history of TBI pass PVTs when completing neuropsychological evaluations (Bigler, 2014). Nevertheless, high PVT failure rates still exist, particularly in those with TBIs, severe cognitive impairment, and dementia (McWhirter et al., 2020). These failure rates continue to exist despite the absence of external gain or evidence of suboptimal effort. For example, in a literature review on PVT failure rates in those with mild TBIs, McWhirter and colleagues (2020) found failure rates ranging from less than 20% to over 50% across various PVTs. Zenisek and colleagues (2016) found that 29.7% of non-litigating patients presenting with dementia and other neurological disorders failed the RDS. Teichner and Wagner (2004) conducted a study on the use of the TOMM with dementia patients and found high rates of misclassification of malingering even when lowering the cut-off score by ten points. Given the substantial evidence in the literature to suggest that PVT failure is somewhat common in neurologically impaired individuals, an important question to consider is: Why do these failures occur and what can be done to mitigate the failures?

The results of the current study suggest that issues may exist in terms of the validity of PVTs for neurologically impaired individuals. These issues may be related to the cognitive abilities associated with various PVTs. For example, the current results suggested that RDS, TOMM, and MSVT Free Recall performance may be more influenced by underlying cognitive constructs (attention and memory), than by motivation or true effort put forth on performance. Substantial evidence exists to suggest that these and other PVTs may be influenced by genuine cognitive impairment rather than issues of effort. RDS has been found to have a significant association with intelligence and attention and a non-significant association with well-known PVTs, suggesting that RDS functions more as a measure of attention rather than as a PVT (Maiman et al., 2019). Furthermore, even when using lower cut-off scores on the RDS, false positive rates of invalid test performance have been found to be 18% among memory impaired individuals (Loring et al., 2016). Similarly, the MSVT FR test has generally been considered as a memory test, rather than as a test of effort (Green, 2004). Armstead-Jehle and Hansen (2016) conducted a factor analysis that confirmed this, as their study demonstrated that the FR subtest of both the MSVT and NV-MSVT loaded onto a memory factor rather than an effort factor. Moreover, the TOMM has been found to have high rates of failure among those with TBIs and dementias, even when utilizing lower cut-off scores. As such, the results of the current study are supported by, and contribute to, the literature to suggest that some PVTs may be heavily influenced by cognitive abilities and may not be valid for those with severe cognitive impairment.

Another validity issue addressed by the current study is whether or not standardized cut-off scores for PVTs should be utilized. Given the high rates of PVT failure in neurologically impaired individuals and the influence their cognitive impairment may have on PVT performance, an argument can be made to suggest that lower cut-off scores should be utilized in order to decrease misclassification of malingering in those with cognitive impairment. This is an argument that has recently received increased attention. Substantial evidence exists to suggest that lowered cut-off scores may be appropriate when evaluating those with TBIs, dementia, MCI, or other forms of cognitive impairment (McWhirter, et al., 2020; Dean et al., 2009; Teichner and Wagner, 2004). This concept has been suggested for various PVTs, including the RDS, MSVT, and the TOMM. This has also been suggested for forced choice tests generally, which is a common paradigm utilized for PVTs. McWhirter and colleagues (2020) argue that a less than chance performance (50% or lower) should be utilized for forced choice tests, rather than the 80%-90% cut-off scores that most forced choice tests utilize, such as the TOMM. They suggest that utilizing higher than chance cut-offs makes the test redundant and susceptible to those with any sort of attentional deficit. The current study adds to the existing literature that PVTs, as traditionally used, may not be valid for those with TBIs, severe cognitive impairment or suspected dementia and that utilizing traditional cut-off scores may lead to the misclassification of malingering.

The issue of the impact of emotional functioning, specifically anxiety and depression, on neuropsychological performance was explored in the current study. This topic has demonstrated mixed results. For example, in a study examining the effects of depression and anxiety on neuropsychological performance in those receiving pre-surgical evaluations, Tsushima and colleagues found age and education to be significant predictors of neuropsychological performance, however depression and anxiety were non-significant (Tsushima et al., 2005).

Similarly, in a study evaluating the impact of depression on neurocognitive performance, Rohling and colleagues found that depression did not impact objective neurocognitive functioning, however depressed patients reported more cognitive problems on self-report measures (Rohling et al., 2002). Other studies, however, have noted cognitive differences between those with and without depression. Mohn and Rund (2016) found those with depression to perform significantly worse on measures of working memory, processing speed, attention, learning, and problem solving as compared to those without depression.

In terms of the impact of emotional functioning on PVT performance, many studies examining the effects of anxiety and depression on PVTs demonstrate that these disorders alone do not commonly result in false positive results (Boone, 2021). Many of the studies, however, excluded those with head injuries, dementia, or other forms of neurological disorders. As such, little research has been conducted on the impact of comorbid anxiety/depression and cognitive impairment on PVT performance. It has been suggested, however, that since anxiety and depression are associated with attention difficulties and other cognitive symptoms, it is possible that comorbid anxiety and/or depression may contribute to high PVT failure rates in those with head injuries (Bigler, 2014). The results of the current study suggest that emotional distress may exacerbate poor performance on cognitive tasks or PVTs. The effects of comorbid anxiety and depression is an important issue that needs to be considered when evaluating those with cognitive impairment, however more research is needed to further explore impact of comorbid psychiatric disorders and TBI and its impact on PVTs and neuropsychological assessment.

In addition to considering the validity of PVTs, appropriate cut-off scores to use, and the impact of emotional distress, other considerations may be taken when utilizing PVTs to assess for test validity and potential feigning. The Slick, Sherman, and Iverson (1999) criteria is the most widely utilized model for assessing the malingering of cognitive deficits (Sherman, Slick, Iverson, 2020). These criteria emphasize the use of multiple sources of information, including neuropsychological testing, PVTs, self-report information, observed behaviors, background history, and collateral informants to identify discrepancies in the data (Slick et al., 2009). This information is then used to determine if an individual may exhibit possible, probable, or definite malingering. In addition, the criteria states that an individual must also exhibit an external incentive and that their presentation should not be fully accounted for by neurological, developmental, or psychiatric factors (Slick et al., 2009). The criteria were recently updated to address other forms of malingering, address issues related to PVTs and SVTs, better outline the meaning and role of inconsistencies in data and to provide exclusionary criteria based on recent malingering research (Sherman et al., 2020). These updated criteria may be used to better assist in evaluating cognitive impairment and potential malingering of deficits.

Additionally, The American Academy of Clinical Neuropsychology (AACN) recently revised their 2009 consensus statement on validity assessment, which updated definitions of validity assessment, malingering and related issues and included latest recommendations regarding validity testing based on current research (Sweet et al., 2021). It includes information on differential diagnoses, methods of evaluating performance validity and symptom validity, clinical considerations when conducting evaluations, ways of conducting thorough neuropsychological evaluations, and research and statistical issues. This updated consensus may also be used to inform neuropsychological and validity assessment.

The current study focused on a unique population: retired professional athletes, and more specifically, retired NFL players. This population is important to consider when discussing head injuries, as football has one of the highest rates of sports-related head injuries, with a concussion occurring in the NFL about once every other game (Clark & Guskiewicz, 2016). Common complaints of those with sports-related concussions include difficulties with attention, memory, and processing speed; fatigue; emotional difficulties, including anxiety, depression, and irritability; sleep difficulties; and physical symptoms such as headaches, nausea, dizziness, and sensory sensitivity (Johnson, Kegal, & Collins, 2011). Not only do NFL players display high rates of sports-related head injuries, they are also susceptible to sustaining repeated head injuries. Increased evidence has demonstrated the cumulative effects of repeated head injuries. Those with multiple head injuries have been found to demonstrate higher cognitive impairment, depression, executive functioning deficits, processing speed deficits, and memory difficulties, as well as worse performance on neuropsychological testing (Johnson et al., 2011; Moser et al., 2007). Furthermore, those with repeated concussions or sub-concussive head injuries are at a higher risk for developing Chronic Traumatic Encephalopathy, or CTE (VanItallie, 2019). CTE is associated with various emotional, behavioral, and cognitive symptoms. It is considered a form of neurodegeneration and is associated with the development of dementia (Gavett et al., 2011).

While repeated head injuries and CTE are associated with severe cognitive impairment, little research has been devoted to validity testing with these populations, as much of the research on validity testing with TBI patients focuses on single TBI events or does not specify the number of head injuries. This yields several concerns regarding validity and cognitive testing with those who sustain repeated head injuries and/or display presentations consistent with CTE. One important consideration is the type of normative sample that is used as a comparison group when interpreting neuropsychological data. While those with sports-related concussions have been found to display cognitive difficulties post-injury, the majority of individuals typically fully recover within one month after the injury (Moser et al., 2007). This is not necessarily the case for those with repeated concussions, as long-term cognitive deficits have been found in those with two or more head injuries (Moser et al., 2007; Johnson et al., 2011). As such, comparing those with one concussion to those who have sustained multiple concussions may impact the validity of cognitive assessments, which in turn may influence the way clinicians interpret data and the types of cognitive profiles they expect to find. A similar issue arises when evaluating those with a CTE presentation, as the cognitive profile may present more similarly to that of a dementia profile, rather than a TBI profile (Gavett et al., 2011; Asken et al., 2016). This is very important to consider when utilizing PVTs with this population, so as to not misinterpret poor performance on PVTs as suboptimal effort or malingering. More research is needed on cognitive evaluations and the use of PVTs in those with repeated concussions and/or potential CTE.

Overall, the current study provided important information regarding the relationship between PVTs and memory performance and the use of PVTs with neurologically impaired individuals. It explored clinically relevant issues regarding performance validity testing, including failure rates of PVTs, validity issues surrounding the use of PVTs for those with severe cognitive impairment, issues surrounding standardized PVT cut-off scores, and the impact of emotional distress on test performance. The findings and implications of the current study were consistent with the current literature on performance validity testing, which suggests that some commonly used PVTs, such as

MSVT, TOMM, and RDS may be heavily influenced by cognitive functioning and that failure on PVT may be a function of genuine impairment rather than suboptimal effort or intentional feigning. The current findings are also consistent with recent literature on performance validity testing suggesting that traditional cut-off scores typically used for PVTs may not be appropriate for those with severe impairment and that lowered cut-off scores may be more appropriate. Finally, the current results addressed issues related to emotional distress and its potential impact on testing performance.

In addition, this study focused on a unique clinical population, retired NFL players. This population represents a large proportion of those with sports-related head injuries and multiple head injuries, however a limited amount of research has been conducted on this unique population. Conducting research on this population can help contribute to our understanding of the assessment and treatment of those with sports-related head injuries and multiple head injuries. As such, the findings from the current study contribute to the growing research on sports-related head injuries, the use of PVTs for this population, and the relationship between PVTs and cognitive performance. The current results provide clinically meaningful information and should be considered when interpreting PVTs as part of this clinical battery for retired NFL players. Further research is needed to: validate these results; further explore the relationship between PVTs and cognitive functioning and the meaning of this relationship; and extend our knowledge and understanding of utilizing PVTs with those with severe cognitive impairment, such as those with multiple head injuries and potential CTE. By conducting this research, clinicians can improve the validity and accuracy of their interpretations of neuropsychological data and use it to better inform assessment and treatment of these unique populations.

Limitations

While the current study demonstrated a predictive ability of PVTs for memory performance and provided important information on the relationship between PVTs and memory performance in a sample of retired NFL players, there were several limitations to the present study. One limitation to the current study is related to the archival nature of the data that was utilized. While the use of archival data provides several benefits, including that it allows researchers to have access to data that they may otherwise not be able to access, reduces costs and time required for data collection, provides retroactive and longitudinal data, and includes data that is typically ready to use for SPSS or SAS analysis, using archival data also presents with several disadvantages (Shultz et al., 2005). Using archival data significantly limits control over data entry and makes it difficult to assess whether the information in the database is accurate. This increases the risk for mistakes or inaccuracies while creating the database, which in turn may impact study findings. This was observed in the current study, as when initially examining the data for Logical Memory II and MMPI-2-RF RC2, one data point from each variable likely reflected a data entry error because it fell outside of possible values. This issue was addressed, however, as the data points subsequently had to be removed. The use of archival datasets also limits what studies can be conducted using a given dataset, as researchers are limited to the variables provided in the dataset and do not have the liberty to collect data on specific areas of interest.

Lack of prior research on this topic also served as a limitation to the present study. While there is substantial research suggesting that memory is impacted by TBIs and that cognitive impairment may impact performance on PVTs, there is a gap in the current literature exploring the predictive ability of PVTs for memory performance. Many studies on PVTs focus on specificity and sensitivity rates, such that they evaluate whether PVTs accurately identify those who may or may not be displaying suboptimal effort. Some studies have evaluated and found a relationship between performance on PVTs and neuropsychological testing, suggesting that as PVT scores decrease, so do scores on neuropsychological tests (Lindem et al., 2003; Perna & Loughan, 2014; Constantinou et al., 2004). Additionally, Green (2007) found that in a TBI sample, the Word Memory Test accounted for 34% of the variance of memory tests and suggested that memory performance was influenced more by effort measures than the effects of brain injury. Moreover, Armistead-Jehle and Hansen (2016) found that MSVT variables accounted for up to 21% of the variance of scores on the RBANS Story Memory subtest. While some research has been conducted to evaluate the relationship between PVTs and cognitive performance, these studies typically focus on just one PVT, rather than the effects of multiple PVTs, such as in the current study. Additionally, different statistical methods were utilized, including examining correlations and univariate tests, and few studies focused on the relationship between PVTs and the WMS subtests. Given the limited prior research focusing on the current topic, an exploratory approach was utilized for the present study, rather than an explanatory approach. This research, however, can be used to lay the foundation for future research exploring the relationship between PVTs and memory performance, as measured by the WMS.

Another limitation to the current study is that TBI-specific factors, such as the number of head injuries sustained by each participant and TBI severity, may have interfered with the current study. These factors were not included in the present study because information on these factors were not provided in the dataset that was utilized. Additionally, the impact of number of head injuries and differences among TBI severity levels fell outside of the scope of the current study, as the focus was on the relationship between PVTs and memory performance. Nevertheless, specific TBI factors may have influenced the current results. Substantial evidence suggests that differences in symptom severity and outcomes exist between those who sustained a single TBI versus those who sustain two or more TBIs. This is such that those who sustain multiple TBIs have been found to display higher levels of cognitive impairment and depression, as well as worsened performance on neuropsychological testing (Johnson et al., 2011; Moser et al., 2007). Those with multiple TBIs are also more likely to develop MCI and CTE (Randolph et al., 2013; VanItallie, 2019). Similarly, differences in symptom severity and outcomes exist between those with differing levels of TBI severity, as concussion symptoms generally tend to be less severe and more transient than moderate/severe TBI symptoms (Moser et al., 2007). While outside of the scope of the current study, these TBI factors may have provided better insight into potential mediating factors of the relationship between PVTs and memory performance. Future research is needed to assess the impact of TBI-specific factors and their impact on PVT and memory performance.

Another limitation to the current study is the methods utilized to explore the impact of anxiety and depression. Comorbid psychiatric issues, namely anxiety and depression, were a topic of interest and were discussed in the current study. However, given the archival nature of the data and the limited information provided on specific comorbid psychiatric diagnoses for each participant, these comorbid disorders could not be directly examined in the current study. Instead, because each participant completed the

MMPI-2-RF, anxiety and depression were explored utilizing MMPI-2-RF RC2, RC7, and RCd scales, as these scales are significantly associated with anxiety and depressive disorders (Ben-Porath, 2012). While these scales are associated with anxiety and depression, utilizing these scales to represent these disorders serves as a limitation, as clinically significant scores on these scales may not necessarily represent anxiety and depression. For example, the RC2 scale has been associated with schizophrenia and PTSD, in addition to depression (Ben-Porath, 2012; Sellbom, 2017). Additionally, relying solely on self-report measures, such as the MMPI-2-RF, to diagnose psychiatric disorders may misrepresent an individual's presentation as testing profiles may be biased or invalid, such as if an individual over-reports or under-reports symptoms. As such, discussion of comorbid anxiety and depression in the current study may be considered theoretical, as explicit comorbid diagnoses could not be included in the study. Future research is needed on the impact of comorbid TBI and anxiety/depression on testing performance in order to validate the current results.

Finally, other comorbid issues may have interfered with the current study. This includes the influence of physical injuries and/or pain. These are important factors to consider, as they may impact neuropsychological testing in several ways. For example, physical injuries that influence or limit motor coordination, fine motor speed, or hand-eye coordination may impact performance on assessments that rely on these abilities. In terms of pain, the development of chronic pain has been associated with increased emotional distress in retired NFL players (Gaetz, 2017). This may in turn increase the risk for developing psychiatric disorders and may impact performance on both memory tests and PVTs. Pain may make it difficult to attend to stimuli during testing and may instead cause an individual to focus attention on the pain (McWhirter et al., 2020). Higher levels of pain have been associated with lower mental processing speed, and those with chronic pain have been found to display impaired neuropsychological performance on immediate memory, delayed memory, language, mental flexibility, and motor functioning as compared to those without chronic pain (Pulles & Oosterman, 2011; Weiner et al., 2006). Moreover, high PVT failure rates have also been found in those presenting with pain-related disorders or concerns. For example, it has been estimated that the malingering rates of those in personal injury, disability, criminal and medical litigation cases is about 30% and about 7-12% in those in non-litigation or compensation seeking cases (Boone, 2021). Furthermore, in a study on those seeking workers' compensation, disability, or personal injury evaluations, about 43% of individuals failed at least one PVT. As such, while the influence of pain and physical injuries were outside of the scope of the current study, they are important factor that may have impacted and interfered with the findings observed in this study.

In spite of the limitations to the present study, the results appear to be valid. The current study was exploratory in nature and provided an initial investigation on the relationship between PVTs and memory performance using the current standardized battery to evaluate retired NFL players. Given this exploratory approach, the intention of this study was to lay the groundwork for future research on the relationship between PVTs and cognitive performance and to contribute to the growing literature on sports-related head injuries among retired professional athletes. Future research is needed to further evaluate the validity and reliability of the present study. Should future studies account for these limitations, stronger evidence may be available to support the current findings and to provide a better understanding of the relationship between PVTs and memory performance. This research can be used to improve our understanding of PVTs, cognitive performance, and sports-related head injuries, which in turn may improve our assessment, treatment, and the quality of care provided to those with sports-related head injuries.

Future Directions

To date, very little research has been conducted on retired NFL players, even though this population represents a large proportion of those with sports-related head injuries and multiple head injuries. Much of the existing research on sports-related head injuries focuses on high school and college level athletes. Additionally, the current sports-related head injury literature focuses largely on the underreporting of symptoms, rather than on performance validity testing and/or neuropsychological performance for this population. Furthermore, while sports organizations, including the NFL, have increasingly made efforts to properly manage and address sports-related head injuries, many former players retired prior to the implementation of these efforts. As a result, retired NFL players possess unique qualities and challenges from those who are currently playing in the NFL.

These gaps in the current sports-related head injury literature makes it difficult to generalize existing research to retired NFL players and also makes it difficult to generalize the results of the current study to other TBI populations. As such, further research focused on NFL players is needed, particularly research focused on neuropsychological performance, performance validity testing and cognitive outcomes, rather than on the underreporting of symptoms. Conducting research on this population can help contribute to our understanding of sports-related head injuries, improve the assessment and treatment of those with sports-related head injuries, and can contribute to the growing TBI literature. Furthermore, given that the battery used in the current study is a standardized battery put forth by the NFL concussion settlement program, future research on this population using this battery can help validate the current results and can provide clinicians with important information regarding decisions made by the NFL concussion settlement program and can help clinicians make better decisions regarding client functioning and treatment.

Additionally, the gaps in the current sports-related head injury literature makes it difficult to generalize existing research to other types of athletes, such as athletes who play in other sports besides football, athletes who speak different languages, or athletes in other countries. Athletes' characteristics and perspectives may vary vastly depending on the type of sport they play and their unique cultural backgrounds, which may in turn impact neuropsychological testing and performance on PVTs. By conducting studies on different athlete populations, researchers may evaluate the validity of PVTs and neuropsychological tests for these different athlete populations and understand how these unique factors may impact neuropsychological performance.

Further research is also needed on the unique neuropsychological profiles of those who sustain multiple TBIs. Individuals with multiple TBIs have been found to be five times more likely to develop MCI and three times more likely to have subjective memory complaints (Randolph et al., 2013). They are also more likely to develop CTE and have higher rates of mood disorders, such as depression. While substantial evidence exists to suggest this, limited research has been conducted on those with multiple head injuries, as many TBI studies focus on single TBI events or do not specify the number of head injuries. Research on repeated head injuries can help clinicians better understand the cognitive, emotional, and behavioral impact of multiple head injuries, as well as the relationship between head injuries and severe cognitive impairment, including MCI, CTE, and dementia. Research on repeated head injuries can also help improve our understanding of TBIs overall, which in turn can improve our assessment and treatment of TBIs.

In terms of research on PVTs, research on this topic is often limited regarding those with dementia, as dementia patients are often excluded from PVTs validation studies (Zenisek et al., 2016). This makes conducting evaluations with dementia patients difficult, as commonly used PVTs may not be valid for dementia populations. This has been commonly suggested in the current PVT literature, as many studies suggest that those with dementia have high PVT failure rates, despite the absence of external gain or evidence of suboptimal effort. This may make it difficult to discern genuine cognitive impairment from suboptimal effort when evaluating those with dementia. As such, future research should focus on the validity of PVTs for dementia patients. This may include conducting validation studies on commonly used PVTs using various dementia populations. This may also include establishing and validating different cut-off scores than traditionally used for PVTs. Future research may also focus on establishing best practice guidelines for utilizing PVTs for dementia populations and reducing the misinterpretation of genuine impairment as invalid performance for this population. By conducting this research, clinicians may improve the accuracy and validity of their neuropsychological evaluations and clinical interpretations.

The current study focused specifically on one cognitive domain: memory. Memory was chosen as a focus for this study as memory deficits are the most commonly reported cognitive deficits after a sports-related concussion and is the most widely studied cognitive domain in TBI patients (Randolph et al., 2005; Vakil, 2005). Other cognitive domains, however, are also commonly impacted by sports-related head injuries, including attention and concentration, processing speed, and executive functioning (Merritt et al., 2017). In order to better understand the relationship between performance on PVTs and cognitive performance, other cognitive domains must be explored. Given that PVTs exist across multiple cognitive domains, including intelligence, memory, processing speed, attention, visual-spatial, language, executive functioning, and motor skills, the relationship between PVTs and cognitive performance may vary based on the cognitive domain being studied. This may have important clinical applications regarding the use of specific PVTs for cognitive performance based on the cognitive domain being evaluated and the cognitive ability associated with a given PVT. Further research is needed to evaluate the relationship between PVTs and cognitive performance across cognitive domains and to evaluate the clinical implications of these relationships.

Future research should focus on the impact of comorbid psychiatric disorders on neuropsychological performance in those with TBIs. Particular emphasis should be placed on comorbid anxiety and depression, as these psychiatric disorders are common consequences of TBIs. Studies focusing on the impact of comorbid anxiety and depression on neuropsychological performance in TBI patients have produced conflicting results, as some studies have found worsened neuropsychological performance in TBI patients with comorbid anxiety and depression compared to TBI patients without comorbid psychiatric disorders, while other studies have found no impact of comorbid anxiety and depression on neuropsychological performance. Many patients with comorbid anxiety and depression, nevertheless, report worsened cognitive deficits, including deficits in memory, attention, processing speed, and executive functioning (Perini et al., 2019). As such, further research is needed to evaluate the impact of comorbid anxiety/depression on cognitive functioning in those with TBI patients. Further research can help clinicians understand the potential impact as well as factors that may mediate this impact.

Finally, future research should focus on the impact of comorbid psychiatric disorders on performance validity testing in those with TBIs. Many studies examining the effects of anxiety and depression on PVTs demonstrate that these psychiatric disorders alone do not commonly result in false positive results (Boone, 2021). Little research, however, has been conducted on the effects of comorbid TBI and psychiatric disorders on PVTs. Given that many TBI patients fail PVTs and that many TBI patients also present with comorbid anxiety and/or depression, exploring the impact of anxiety/depression on performance validity testing in TBI patients may contribute to our understanding of utilizing PVTs with this population and in turn improve our neuropsychological assessments of TBI patients.

Overall, future research focused on sports-related head injuries is needed, particularly focused on retired professional athletes, such as the retired NFL players examined in the current study. Future research is needed focusing on PVTs and neuropsychological testing of those with sports-related head injuries, as this population represents a large proportion of TBI patients. Additionally, further research is needed to evaluate the impact of repeated head injuries on PVTs and cognitive functioning. The validity of PVTs for those with dementia also needs further exploration. Finally, future research in this area should focus on the relationship between PVTs and cognitive performance across cognitive domains, as well as the impact of comorbid anxiety/depression on cognitive functioning and PVTs in those with TBI patients. By conducting further research on these topics, researchers and clinicians can work to improve our understanding, assessment and treatment of those with sports-related head injuries, as well as improve the overall quality of care provided to these patients. Moreover, further research on these topics may also pave the way for improvements policy making regarding sports-related head injuries within the NFL and other sports organizations. Through increased efforts towards understanding sports-related head injuries, we may further develop the measures taken to prevent sports-related head injuries, improve the validity of assessments following a head injury, and increase advocacy for athlete care and wellbeing.

References

Amen, D. G., Newberg, A., Thatcher, R., Jin, Y., Wu, J., Keator, D., & Willeumier, K. (2011). Impact of playing American professional football on long-term brain function. *The Journal of Neuropsychiatry and Clinical Neurosciences, 23*(1), 98-106. http://search.proquest.com.ezproxylocal.library.nova.edu/docview/853645630?accountid=6579

Armistead-Jehle, P., & Buican, B. (2013). Comparison of select advanced clinical solutions embedded effort measures to the word memory test in the detection of suboptimal effort. *Archives of Clinical Neuropsychology: The Official Journal of the National Academy of Neuropsychologists, 28*(3), 297-301. doi:http://dx.doi.org.ezproxylocal.library.nova.edu/10.1093/arclin/act017

Armistead-Jehle, P., & Hansen, C. L. (2016). Factor analysis of the MSVT, NV-MSVT, and RBANS memory subtests. *Archives of Clinical Neuropsychology: The Official Journal of the National Academy of Neuropsychologists, 31*(5), 465-471. doi:http://dx.doi.org.ezproxylocal.library.nova.edu/10.1093/arclin/acw033

Asken, B. M., Sullan, M. J., Snyder, A. R., Houck, Z. M., Bryant, V. E., Hizel, L. P., McLaren, M. E., Dede, D. E., Jaffee, M. S., DeKosky, S. T., & Bauer, R. M. (2016). Factors Influencing Clinical Correlates of Chronic Traumatic Encephalopathy (CTE): A Review. *Neuropsychology Review*, *26*(4), 340–363. https://doi-org.ezproxylocal.library.nova.edu/10.1007/s11065-016-9327-z

Bain, K. M., & Soble, J. R. (2019). Validation of the advanced clinical solutions word choice test (WCT) in a mixed clinical sample: Establishing classification accuracy, Sensitivity/Specificity, and cutoff scores. *Assessment, 26(7),* 1320-1328. doi:http://dx.doi.org.ezproxylocal.library.nova.edu/10.1177/1073191117725172

Bashem, J. R., Rapport, L. J., Miller, J. B., Hanks, R. A., Axelrod, B. N., & Millis, S. R. (2014). Comparisons of five performance validity indices in bona fide and simulated traumatic brain injury. *The Clinical neuropsychologist*, *28*(5), 851–875. https://doi.org/10.1080/13854046.2014.927927

Ben-Porath, Y. S. (2012). *Interpreting the MMPI-2-RF.* University of Minnesota Press.

Bigler E. D. (2014). Effort, symptom validity testing, performance validity testing and traumatic brain injury. *Brain Injury*, *28*(13-14), 1623–1638. https://doi.org/10.3109/02699052.2014.947627

Boone, K. B. (Ed.). (2021). *Assessment of feigned cognitive impairment: A neuropsychological perspective* (2nd ed.). The Guilford Press.

Carlozzi, N. E., Grech, J., & Tulsky, D. S. (2013). Memory functioning in individuals with traumatic brain injury: An examination of the Wechsler Memory Scale–Fourth edition (WMS–IV). *Journal of Clinical and Experimental Neuropsychology, 35*(9), 906-914. doi:http://dx.doi.org.ezproxylocal.library.nova.edu/10.1080/13803395.2013.833

Chase, D., Schatz, P., Smyk, N., & Franks, R. R. (2018). The stability of engagement over comprehensive neuropsychological assessment in student athletes diagnosed with sports related concussion. *Developmental Neuropsychology, 43*(4), 345-355. doi:http://dx.doi.org.ezproxylocal.library.nova.edu/10.1080/87565641.2018.1428326

Clark, M., & Guskiewicz K. (2016). Sport-Related Traumatic Brain Injury. In D. Laskowitz & G. Grant (Eds.), *Translational Research in Traumatic Brain Injury*. CRC Press/Taylor and Francis Group. https://www.ncbi.nlm.nih.gov/books/NBK326721/

Center for Disease Control and Prevention. (2020, January 8). *Traumatic Brain Injury Prevention.* https://www.cdc.gov/injury/stateprograms/topic_traumatic-brain-injury.html

Chun, M. M., & Turk-Browne, N. B. (2007). Interactions between attention and memory. *Current Opinion in Neurobiology, 17*(2), 177-184. https://doi.org/10.1016/j.conb.2007.03.005

Coughlin, J. M., Wang, Y., Munro, C. A., Ma, S., Yue, C., Chen, S., Airan, R., Kim, P. K., Adams, A. V., Garcia, C., Higgs, C., Sair, H. I., Sawa, A., Smith, G., Lyketsos, C. G., Caffo, B., Kassiou, M., Guilarte, T. R., & Pomper, M. G. (2015). Neuroinflammation and brain atrophy in former NFL players: An in vivo multimodal imaging pilot study. *Neurobiology of disease*, *74*, 58–65. https://doi.org/10.1016/j.nbd.2014.10.019

Constantinou, M., Bauer, L., Ashendorf, L., Fisher, J. M., & McCaffrey, R. J. (2005). Is poor performance on recognition memory effort measures indicative of generalized poor performance on neuropsychological tests?. *Archives of Clinical Neuropsychology: The Official Journal of the National Academy of Neuropsychologists*, *20*(2), 191–198. https://doi.org/10.1016/j.acn.2004.06.002

Covassin, T., Savage, J. L., Bretzin, A. C., & Fox, M. E. (2018). Sex differences in sport-related concussion long-term outcomes. *International Journal of Psychophysiology*, *132*, 9–13. https://doi-org.ezproxylocal.library.nova.edu/10.1016/j.ijpsycho.2017.09.010

Cowan N. (2008). What are the differences between long-term, short-term, and working memory?. *Progress in Brain Research*, *169*, 323–338. https://doi.org/10.1016/S0079-6123(07)00020-9

Cubon, V. A., Putukian, M., Boyer, C., & Dettwiler, A. (2011). A diffusion tensor imaging study on the white matter skeleton in individuals with sports-related concussion. *Journal of Neurotrauma, 28*(2), 189-201. doi:http://dx.doi.org.ezproxylocal.library.nova.edu/10.1089/neu.2010.1430

Dandachi-FitzGerald, B., & Merckelbach, H. (2013). Feigning ≠ Feigning a Memory Deficit: The Medical Symptom Validity Test as an Example. *Journal of Experimental Psychopathology*, 46–63. https://doi.org/10.5127/jep.025511

Delain, S. L., Stafford, K. P., & Ben-Porath, Y. (2003). Use of the TOMM in a criminal court forensic assessment setting. *Assessment, 10*(4), 370-381. doi:http://dx.doi.org.ezproxylocal.library.nova.edu/10.1177/1073191103259156

Dean, A. C., Victor, T. L., Boone, K. B., Philpott, L. M., & Hess, R. A. (2009). Dementia and effort test performance. *The Clinical Neuropsychologist, 23*(1), 133-152. https://doi.org/10.1080/13854040701819050

Didehbani, N., Munro Cullum, C., Mansinghani, S., Conover, H., & Hart, J. (2013). Depressive symptoms and concussions in aging retired NFL players. *Archives of Clinical Neuropsychology: The Official Journal of the National Academy of Neuropsychologists, 28*(5), 418-424. doi:http://dx.doi.org.ezproxylocal.library.nova.edu/10.1093/arclin/act028

Donders, J. (2005). Performance on the test of memory malingering in a mixed pediatric sample. *Child Neuropsychology, 11*(2), 221-227. doi:http://dx.doi.org.ezproxylocal.library.nova.edu/10.1080/09297040490917298

Duncan, A. (2005). The impact of cognitive and psychiatric impairment of psychotic disorders on the test of memory malingering (TOMM). *Assessment, 12*(2), 123-129. doi:http://dx.doi.org.ezproxylocal.library.nova.edu/10.1177/1073191105275512

Dziemianowicz, M. S., Kirschen, M. P., Pukenas, B. A., Laudano, E., Balcer, L. J., & Galetta, S. L. (2012). Sports-related concussion testing. *Current Neurology and Neuroscience Reports, 12*(5), 547-59. doi:http://dx.doi.org.ezproxylocal.library.nova.edu/10.1007/s11910-012-0299-y

Echemendia, R. J., Thelen, J., Meeuwisse, W., Comper, P., Hutchison, M. G., & Bruce, J. M. (2020). Testing the hybrid battery approach to evaluating sports-related concussion in the National Hockey League: A factor analytic study. *The Clinical Neuropsychologist*, *34*(5), 899–918. https://doi-org.ezproxylocal.library.nova.edu/10.1080/13854046.2019.1690051

Erdal, K. (2012). Neuropsychological testing for sports-related concussion: How athletes can sandbag their baseline testing without detection. *Archives of Clinical Neuropsychology: The Official Journal of the National Academy of Neuropsychologists, 27*(5), 473-479. doi:http://dx.doi.org.ezproxylocal.library.nova.edu/10.1093/arclin/acs050

Faul, M., & Coronado, V. (2015). Epidemiology of traumatic brain injury. *Handbook of Clinical Neurology, 127*, 3-13. doi:http://dx.doi.org.ezproxylocal.library.nova.edu/10.1016/B978-0-444-52892-6.00001-5

Field, A. (2009). *Discovering statistics using IBM SPSS statistics* (3rd ed.). SAGE Publications.

Fields, L., Didehbani, N., Hart, J., & Cullum, C. M. (2020). No Linear Association Between Number of Concussions or Years Played and Cognitive Outcomes in Retired NFL Players. *Archives of Clinical Neuropsychology: The Official Journal of the National Academy of Neuropsychologists*, *35*(3), 233–239. https://doi.org/10.1093/arclin/acz008

Forbey, J. D., Lee, T. T. C., & Handel, R. W. (2010). Correlates of the MMPI-2-RF in a college setting. *Psychological Assessment, 22*(4), 737-744. doi:http://dx.doi.org.ezproxylocal.library.nova.edu/10.1037/a0020645

Friedland, D., & Hutchinson, P. (2013). Classification of traumatic brain injury. *Advances in Clinical Neuroscience and Rehabilitation.* http://www.acnr.co.uk/wp-content/uploads/2013/07/ACNRJA13_rehab1.pdf

Frost, R. B., Farrer, T. J., Primosch, M., & Hedges, D. W. (2013). Prevalence of traumatic brain injury in the general adult population: A meta-analysis. *Neuroepidemiology, 40*(3), 154-159. doi:http://dx.doi.org.ezproxylocal.library.nova.edu/10.1159/000343275

Gaetz M. (2017). The multi-factorial origins of Chronic Traumatic Encephalopathy (CTE) symptomology in post-career athletes: The athlete post-career adjustment (AP-CA) model. *Medical Hypotheses*, *102*, 130–143. https://doi-org.ezproxylocal.library.nova.edu/10.1016/j.mehy.2017.03.023

Gardner, A., Kay-Lambkin, F., Stanwell, P., Donnelly, J., Williams, W. H., Hiles, A., Schofield, P., Levi, C., & Jones, D. K. (2012). A systematic review of diffusion tensor imaging findings in sports-related concussion. *Journal of Neurotrauma*, *29*(16), 2521–2538. https://doi.org/10.1089/neu.2012.2628

Gass, C. S. (1991). Emotional variables and neuropsychological test performance. *Journal of Clinical Psychology, 47*(1), 100. http://search.proquest.com.ezproxylocal.library.nova.edu/docview/236934836?accountid=6579

Gavett, B. E., Stern, R. A., & McKee, A. C. (2011). Chronic traumatic encephalopathy: a potential late effect of sport-related concussive and subconcussive head trauma. *Clinics in Sports Medicine*, *30*(1), 179–xi. https://doi.org/10.1016/j.csm.2010.09.007

Green, P. (2004) *Manual for the Medical Symptom Validity Test for Windows.* Green's Publishing.

Green, P. (2007). The pervasive influence of effort on neuropsychological tests. *Physical Medicine and Rehabilitation Clinics of North America*, *18*(1), 43–vi. https://doi.org/10.1016/j.pmr.2006.11.002

Green, P., Montilo, J., & Brockhaus, R. (2011). High specificity of the word memory test and medical symptom validity test in groups with severe verbal memory impairment. *Applied Neuropsychology, 18*(2), 86-94. doi: 10.1080/09084282.2010.523389.

Greher, M. R., & Wodushek, T. R. (2017). Performance validity testing in neuropsychology: Scientific basis and clinical application-A brief review. *Journal of Psychiatric Practice, 23*(2), 134-140. doi:http://dx.doi.org.ezproxylocal.library.nova.edu/10.1097/PRA.0000000000000218

Jasinski, L. J., Berry, D. T., Shandera, A., L., & Clark, J. A. (2011). Use of the Weschler adult intelligence scale digit span subtest for malingering detection: A meta-analytic review. *Journal of Clinical and Experimental Neuropsychology, 33*(3), 300-314. doi:10.1080/13803395.2010.516743.

Jelicic, M., Ceunen, E., Peters, M. J. V., & Merckelbach, H. (2011). Detecting coached feigning using the test of memory malingering (TOMM) and the structured inventory of malingered symptomatology (SIMS). *Journal of Clinical Psychology, 67*(9), 850-855. doi:http://dx.doi.org.ezproxylocal.library.nova.edu/10.1002/jclp.20805

Johnson, E. W., Kegel, N. E., & Collins, M. W. (2011). Neuropsychological assessment of sport-related concussion. *Clinics in Sports Medicine*, *30*(1), 73–ix. https://doi.org/10.1016/j.csm.2010.08.007

Jordan, B. D., & Zimmerman, R. D. (1988). Magnetic resonance imaging in amateur boxers. *Archives of Neurology*, *45*(11), 1207-1208. https://doi.org/10.1001/archneur.1988.00520350045014

Jurick, S. M., Crocker, L. D., Merritt, V. C., Hoffman, S. N., Keller, A. V., Eglit, G., Thomas, K. R., Norman, S. B., Schiehser, D. M., Rodgers, C. S., Twamley, E. W., & Jak, A. J. (2020). Psychological Symptoms and Rates of Performance Validity Improve Following Trauma-Focused Treatment in Veterans with PTSD and History of Mild-to-Moderate TBI. *Journal of the International Neuropsychological Society: JINS*, *26*(1), 108–118. https://doi.org/10.1017/S1355617719000997

Kang, H., Zhao, F., You, L., Giorgetta, C., D, V., Sarkhel, S., & Prakash, R. (2014). Pseudo-dementia: A neuropsychological review. *Annals of Indian Academy of Neurology*, *17*(2), 147–154. https://doi.org/10.4103/0972-2327.132613

Kellogg, R. T. (1997). *Cognitive Psychology*. Sage Publications.

Kelly, A. B., Zimmerman, R. D., Snow, R. B., Gandy, S. E., Heier, L. A., & Deck, M. D. (1988). Head trauma: Comparison of MR and CT--experience in 100 patients. *AJNR.American Journal of Neuroradiology, 9*(4), 699-708.

http://search.proquest.com.ezproxylocal.library.nova.edu/docview/78329266?accountid=6579

Kerr, Z. Y., Register-Mihalik, J. K., Kay, M. C., DeFreese, J. D., Marshall, S. W., & Guskiewicz, K. M. (2018). Concussion Nondisclosure During Professional Career Among a Cohort of Former National Football League Athletes. *The American Journal of Sports Medicine*, *46*(1), 22–29. https://doi.org/10.1177/0363546517728264

Kirkwood, M. W., Hargrave, D. D., & Kirk, J. W. (2011). The value of the WISC-IV digit span subtest in detecting noncredible performance during pediatric neuropsychological examinations. *Archives of Clinical Neuropsychology, 26*(5), 377-384. doi:http://dx.doi.org.ezproxylocal.library.nova.edu/10.1093/arclin/acr040

Kiewel, N. A., Wisdom, N. M., Bradshaw, M. R., Pastorek, N. J., & Strutt, A. M. (2012). A Retrospective Review of Digit Span-Related Effort Indicators in Probable Alzheimer's Disease Patients. *The Clinical Neuropsychologist, 26*(6), 965-974. DOI: 10.1080/13854046.2012.694478

Langlois, J. A., Rutland-Brown, W., & Wald, M. M. (2006). The epidemiology and impact of traumatic brain injury: A brief overview. *The Journal of Head Trauma Rehabilitation, 21*(5), 375-378. doi:http://dx.doi.org.ezproxylocal.library.nova.edu/10.1097/00001199-200609000-00001

Lezak, M.D., Howieson, D.B., Bigler, E.D., & Tranel, D. (2012). *Neuropsychological Assessment*. (5th ed.). Oxford University Press.

Lindem, K., White, R. F., Heeren, T., Proctor, S. P., Krengel, M., Vasterling, J., Wolfe, J., Sutker, P. B., Kirkley, S., & Keane, T. M. (2003). Neuropsychological performance in gulf war ear veterans: Motivational factors and effort. *Journal of Psychopathology and Behavioral Assessment, 25*(2), 129-138. DOI: 10.1023/A:1023399100404

Loring, D. W., Goldstein, F. C., Chen, C., Drane, D. L., Lah, J. J., Zhao, L., & Larrabee, G. J. (2016). False-Positive Error Rates for Reliable Digit Span and Auditory Verbal Learning Test Performance Validity Measures in Amnestic Mild Cognitive Impairment and Early Alzheimer Disease. *Archives of Clinical Neuropsychology*, *31*(4), 313–331. https://doiorg.ezproxylocal.library.nova.edu/10.1093/arclin/acw014

Lubrini, G., Viejo-Sobera, R., Periáñez, J. A., Cicuendez, M., Castaño, A. M., González-Marqués, J., Lagares, A., & Ríos-Lago, M. (2020). Evolution of cognitive impairment after a traumatic brain injury: is there any improvement after controlling the practice effect?. Evolución de las alteraciones cognitivas tras un traumatismo craneoencefálico: ¿hay mejoría tras controlar el efecto de la práctica?. *Revista de Neurologia*, *70*(2), 37–44. https://doi.org/10.33588/rn.7002.2019233

Maiman, M., Del Bene, V. A., MacAllister, W. S., Sheldon, S., Farrell, E., Arce Rentería, M., Slugh, M., Nadkarni, S. S., & Barr, W. B. (2019). Reliable Digit Span: Does it Adequately Measure Suboptimal Effort in an Adult Epilepsy Population?. *Archives of Clinical Neuropsychology: The Official Journal of the National Academy of Neuropsychologists*, *34*(2), 259–267. https://doi.org/10.1093/arclin/acy027

Maroon, J. C., Winkelman, R., Bost, J., Austin, A., Mathyssek, C., & Miele, V. (2015). Chronic traumatic encephalopathy in contact sports: A systematic review of all reported pathological cases. *PLoS One, 10*(2) doi:http://dx.doi.org.ezproxylocal.library.nova.edu/10.1371/journal.pone.0117338

Marar, M., McIlvain, N. M., Fields, S. K., & Comstock, R. D. (2012). Epidemiology of concussions among united states high school athletes in 20 sports. *The American Journal of Sports Medicine, 40*(4), 747-755. doi:http://dx.doi.org.ezproxylocal.library.nova.edu/10.1177/0363546511435626

McKee, A. C., Stern, R. A., Nowinski, C. J., Stein, T. D., Alvarez, V. E., Daneshvar, D. H., Lee, H. S., Wojtowicz, S. M., Hall, G., Baugh, C. M., Riley, D. O., Kubilus, C. A., Cormier, K. A., Jacobs, M. A., Martin, B. R., Abraham, C. R., Ikezu, T., Reichard, R. R., Wolozin, B. L., Budson, A. E., … Cantu, R. C. (2013). The spectrum of disease in chronic traumatic encephalopathy. *Brain: A Journal of Neurology*, *136*(Pt 1), 43–64. https://doi.org/10.1093/brain/aws307

McWhirter, L., Ritchie, C. W., Stone, J., & Carson, A. (2020). Performance validity test failure in clinical populations—a systematic review. *Journal of Neurology, Neurosurgery and Psychiatry, 91*(9), 945-952. doi:http://dx.doi.org/10.1136/jnnp-2020-323776

Meier, T. B., Brummel, B. J., Singh, R., Nerio, C. J., Polanski, D. W., & Bellgowan, P. S. (2015). The underreporting of self-reported symptoms following sports-related concussion. *Journal of Science and Medicine in Sport, 18*(5), 507-511. DOI: 10.1016/j.jsams.2014.07.008

Menon, D. K., Schwab, K., Wright, D. W., & Maas, A. I. (2010). Position statement: Definition of traumatic brain injury. *Archives of Physical Medicine and Rehabilitation, 91*(11), 1637-1640. doi:http://dx.doi.org.ezproxylocal.library.nova.edu/10.1016/j.apmr.2010.05.017

Merritt, V. C., Meyer, J. E., Cadden, M. H., Roman, C. A., Ukueberuwa, D. M., Shapiro, M. D., & Arnett, P. A. (2017). Normative Data for a Comprehensive Neuropsychological Test Battery used in the Assessment of Sports-Related Concussion. *Archives of Clinical Neuropsychology: The official Journal of the National Academy of Neuropsychologists*, *32*(2), 168–183. https://doi.org/10.1093/arclin/acw090

Miller, J. B., Millis, S. R., Rapport, L. J., Bashem, J. R., Hanks, R. A., & Axelrod, B. N. (2011). Detection of insufficient effort using the advanced clinical solutions for the Wechsler memory scale, fourth edition. *The Clinical Neuropsychologist, 25*(1), 160-172. doi:http://dx.doi.org.ezproxylocal.library.nova.edu/10.1080/13854046.2010.533197

Mittenberg, W., Patton, C., Canyock, E. M., & Condit, D. C. (2002). Base rates of malingering and symptom exaggeration. *Journal of Clinical and Experimental Neuropsychology, 24*(8), 1094-1102. http://search.proquest.com.ezproxylocal.library.nova.edu/docview/72867814?accountid=6579

Mohn, C., & Rund, B. R. (2016). Neurocognitive profile in major depressive disorders: relationship to symptom level and subjective memory complaints. *BMC Psychiatry*, *16*, 108. https://doi.org/10.1186/s12888-016-0815-8

Moradi, E., Hallikainen, I., Hanninen, T., & Tohka, J. (2017). Rey's Auditory Verbal Learning Test scores can be predicted from whole brain MRI in Alzheimer's disease. *Neuroimage: Clinical, 13*, 415-427. https://doi.org/10.1016/j.nicl.2016.12.011

Moser, R. S., Iverson, G. L., Echemendia, R. J., Lovell, M. R., Schatz, P., Webbe, F. M., Ruff, R. M., Barth, J. T., NAN Policy and Planning Committee, & Donna K. Broshek, Shane S. Bush, Sandra P. Koffler, Cecil R. Reynolds, Cheryl H. Silver (2007). Neuropsychological evaluation in the diagnosis and management of sports-related concussion. *Archives of Clinical Neuropsychology: The Official Journal of the National Academy of Neuropsychologists*, *22*(8), 909–916. https://doi.org/10.1016/j.acn.2007.09.004

Murdaugh, D. L., King, T. Z., Sun, B., Jones, R. A., Ono, K. E., Reisner, A., & Burns, T. G. (2018). Longitudinal changes in resting state connectivity and white matter integrity in adolescents with sports-related concussion. *Journal of the International Neuropsychological Society: JINS, 24*(8), 781-792. doi:http://dx.doi.org.ezproxylocal.library.nova.edu/10.1017/S1355617718000413

Nie, J., Zhang, Z., Wang, B., Li, H., Xu, J., Wu, S., Zhu, C., Yang, X., Liu, B., Wu, Y., Tan, S., Wen, Z., Zheng, J., Shu, S., & Ma, L. (2019). Different memory patterns of digits: A functional MRI study. *Journal of Biomedical Science*, *26*(1), 22. https://doi.org/10.1186/s12929-019-0516-y

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. doi:10.1080/13803395.2019.1644294

Omalu, B., Bailes, J., Hamilton, R. L., Kamboh, M. I., Hammers, J., Case, M., & Fitzsimmons, R. (2011). Emerging histomorphologic phenotypes of chronic traumatic encephalopathy in american athletes. *Neurosurgery, 69*(1), 173-83; discussion 183. doi:http://dx.doi.org.ezproxylocal.library.nova.edu/10.1227/NEU.0b013e318212bc7b

Oudman, E., Krooshof, E., van Oort, R., Lloyd, B., Wijnia, J. W., & Postma, A. (2019). Effects of Korsakoff Amnesia on performance and symptom validity testing. *Applied Neuropsychology. Adult*, 1–9. Advance online publication. doi:10.1080/23279095.2019.1576180

Perna, R., & Loughan, A. R. (2014). The influence of effort on neuropsychological performance in children: is performance on the TOMM indicative of neuropsychological ability?. *Applied Neuropsychology. Child*, *3*(1), 31–37. https://doi.org/10.1080/21622965.2012.686339

Perini, G., Cotta Ramusino, M., Sinforiani, E., Bernini, S., Petrachi, R., & Costa, A. (2019). Cognitive impairment in depression: Recent advances and novel treatments. *Neuropsychiatric Disease and Treatment*, *15*, 1249–1258. https://doi.org/10.2147/NDT.S199746

Pituch, K. A., & Stevens, J. P. (2015). *Applied Multivariate Statistics for the Social Sciences: Analyses with SAS and IBM's SPSS* (6th ed.). Routledge.

Pulles, W. L. J. A., & Oosterman, J. M. (2011). The role of neuropsychological performance in the relationship between chronic pain and functional physical impairment. *Pain Medicine*, *12*(12), 1769–1776. https://doi.org/10.1111/j.1526-4637.2011.01266.x

Randolph, C., Karantzoulis, S., & Guskiewicz, K. (2013). Prevalence and characterization of mild cognitive impairment in retired national football league players. *Journal of the International Neuropsychological Society: JINS, 19*(8), 873-880. doi:http://dx.doi.org.ezproxylocal.library.nova.edu/10.1017/S1355617713000805

Randolph, C., McCrea, M., & Barr, W. B. (2005). Is neuropsychological testing useful in the management of sport-related concussion? *Journal of Athletic Training, 40*(3), 139-

154.
http://search.proquest.com.ezproxylocal.library.nova.edu/docview/620944059?accountid=6579

Rapoport, M. J., McCullagh, S., Shammi, P., & Feinstein, A. (2005). Cognitive impairment associated with major depression following mild and moderate traumatic brain injury. *The Journal of Neuropsychiatry and Clinical Neurosciences*, *17*(1), 61–65. https://doi.org/10.1176/jnp.17.1.61

Rohling, M. L., Green, P., Allen III, L. M., & Iverson, G. L. (2002). Depressive symptoms and neurocognitive test scores in patients passing symptom validity tests. *Archives of Clinical Neuropsychology*, *17*(3), 205. https://doi-org.ezproxylocal.library.nova.edu/10.1093/arclin/17.3.205

Romero, I. E., Toorabally, N., Burchett, D., Tarescavage, A. M., & Glassmire, D. M. (2017). Mapping the MMPI–2–RF Substantive Scales onto Internalizing, Externalizing, and Thought Dysfunction Dimensions in a Forensic Inpatient Setting. *Journal of Personality Assessment*, *99*(4), 351–362. https://doi-org.ezproxylocal.library.nova.edu/10.1080/00223891.2016.1223681

Schroeder, R. W., Twumasi-Ankrah, P., Baade, L. E., & Marshall, P. S. (2012). Reliable Digit Span: A Systematic Review and Cross-Validation Study. *Assessment*, *19*(1), 21–30. https://doi.org/10.1177/1073191111428764

Schwarzbold, M., Diaz, A., Martins, E. T., Rufino, A., Amante, L. N., Thais, M. E., Quevedo, J., Hohl, A., Linhares, M. N., & Walz, R. (2008). Psychiatric disorders and traumatic brain injury. *Neuropsychiatric Disease and Treatment*, *4*(4), 797–816. https://doi.org/10.2147/ndt.s2653

Shultz, K.S., Hoffman, C.C., & Reiter-Palmon, R. (2005). Using Archival Data for I-O Research: Advantages, Pitfalls, Sources, and Examples. *The Industrial-Organizational Psychologist, 42*(3), 31-37. DOI:10.1037/e579182011-004

Sellbom, M. (2017). MMPI-2 and MMPI-2-RF Restructured Clinical Scales. In: Zeigler-Hill V., Shackelford T. (eds.). *Encyclopedia of Personality and Individual Differences*. Springer, Cham. https://doi.org/10.1007/978-3-319-28099-8_70-1

Sellbom, M., Waugh, M. H., & Hopwood, C. J. (2018). Development and validation of personality disorder spectra scales for the MMPI-2-RF. *Journal of Personality Assessment, 100*(4), 406-420. https://doi.org/10.1080/00223891.2017.1407327

Sherman, E. M. S., Slick, D. J., & Iverson, G. L. (2020). Multidimensional malingering criteria for neuropsychological assessment: A 20-Year update of the malingered neuropsychological dysfunction criteria. *Archives of Clinical Neuropsychology*, 35(6),735-764, https://doi.org/10.1093/arclin/acaa019

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. https://doi.org/10.1076/1385-4046(199911)13:04;1-Y;FT545

Stamm, J. M., Bourlas, A. P., Baugh, C. M., Fritts, N. G., Daneshvar, D. H., Martin, B. M., McClean, M. D., Tripodis, Y., & Stern, R. A. (2015). Age of first exposure to football and later-life cognitive impairment in former NFL players. *Neurology*, *84*(11), 1114–1120. https://doi.org/10.1212/WNL.0000000000001358

Suesse, M., Wong, V. W., Stamper, L. L., Carpenter, K. N., & Scott, R. B. (2015). Evaluating the Clinical Utility of the Medical Symptom Validity Test (MSVT): A

Clinical Series. *The Clinical Neuropsychologist*, *29*(2), 214–231. https://doi.org/10.1080/13854046.2015.1022226

Sweet, J. J., Heilbronner, R. L., Morgan, J. E., Larrabee, G. J., Rohling, M. L., Boone, K. B., Kirkwood, M. W., Schroeder, R. W., Suhr, J. A., & Conference Participants (2021). American Academy of Clinical Neuropsychology (AACN) 2021 consensus statement on validity assessment: Update of the 2009 AACN consensus conference statement on neuropsychological assessment of effort, response bias, and malingering. *The Clinical Neuropsychologist*, 1–54. Advance online publication. https://doi.org/10.1080/13854046.2021.1896036

Teichner, G., & Wagner, M. T. (2004). Test of Memory Malingering (TOMM): Normative data from cognitively intact, cognitively impaired, and elderly patients with dementia. *Archives of Clinical Neuropsychology, 19*(3), 455–464. https://doi.org/10.1016/S0887-6177(03)00078-7

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

Tombaugh, T. N. (1997). The test of memory malingering (TOMM): Normative data from cognitively intact and cognitively impaired individuals. *Psychological Assessment, 9*(3), 260-268. doi:http://dx.doi.org.ezproxylocal.library.nova.edu/10.1037/1040-3590.9.3.260

Tsushima, W. T., Johnson, D. B., Lee, J. D., Matsukawa, J. M., & Fast, K. M. S. (2005). Depression, anxiety and neuropsychological test scores of candidates for coronary artery bypass graft surgery. *Archives of Clinical Neuropsychology, 20*(5), 667-673. doi:http://dx.doi.org.ezproxylocal.library.nova.edu/10.1016/j.acn.2005.04.003

Vakil, E. (2005). The effect of moderate to severe traumatic brain injury (TBI) on different aspects of memory: A selective review. *Journal of Clinical and Experimental Neuropsychology*, *27*(8), 977–1021. doi:10.1080/13803390490919245

Van Der Heijden, P. T., Egger, J. M., & Derksen, J. L. (2010). Comparability of scores on the MMPI–2–RF scales generated with the MMPI–2 and MMPI–2–RF booklets. *Journal of Personality Assessment, 92*(3)*,* 254-259. DOI: 10.1080/00223891003670208

VanItallie, T.B. (2019). Traumatic brain injury (TBI) in collision sports: Possible mechanisms of transformation into chronic traumatic encephalopathy (CTE). *Metabolism*. doi:10.1016/j.metabol.2019.07.007

Wechsler, D. (2009). *Advanced Clinical Solutions for WAIS-IV and WMS-IV*. Pearson.

Weiner, D. K., Rudy, T. E., Morrow, L., Slaboda, J., & Lieber, S. (2006). The relationship between pain, neuropsychological performance, and physical function in community-dwelling older adults with chronic low back pain. *Pain Medicine, 7*(1), 60–70. https://doi.org/10.1111/j.1526-4637.2006.00091.x

Whitney, K. A., Davis, J. J., Shepard, P. H., Bertram, D. M., & Adams, K. M. (2009). Digit span age scaled score in middle-aged military veterans: is it more closely associated with TOMM failure than reliable digit span?. *Archives of Clinical Neuropsychology: The Official Journal of the National Academy of Neuropsychologists*, *24*(3), 263–272. doi:10.1093/arclin/acp034

Willis, P. F., Farrer, T. J., & Bigler, E. D. (2011). Are Effort Measures Sensitive to Cognitive Impairment?. *Military Medicine, 176*(12), 1426–1431, https://doi.org/10.7205/MILMED-D-11-00168

Wolf, E. J., Miller, M. W., Orazem, R. J., Weierich, M. R., Castillo, D. T., Milford, J., Kaloupek, D. G., & Keane, T. M. (2008). The MMPI-2 restructured clinical scales in the assessment of posttraumatic stress disorder and comorbid disorders. *Psychological Assessment*, *20*(4), 327–340. https://doi.org/10.1037/a0012948

Yengo-Kahn, A., Hale, A. T., Zalneraitis, B. H., Zuckerman, S. L., Sills, A. K., & Solomon, G. S. (2016). The sport concussion assessment tool: A systematic review. *Neurosurgical Focus, 40*(4), 1. doi:http://dx.doi.org.ezproxylocal.library.nova.edu/10.3171/2016.1.FOCUS15611

Zenisek, R., Millis, S. R., Banks, S. J., & Miller, J. B. (2016). Prevalence of below-criterion Reliable Digit Span scores in a clinical sample of older adults. *Archives of Clinical Neuropsychology: The Official Journal of the National Academy of Neuropsychologists*, *31*(5), 426–433. https://doi.org/10.1093/arclin/acw025