9-5-2008

Hopkins Verbal Learning Test - Revised: Standardization Data of Percentage Retention and Comparison of Retention Rates with Logical Memory Subtests of the Wechsler Memory Scales - Revised

Stephen Fink
Indiana University of Pennsylvania

HOPKINS VERBAL LEARNING TEST - REVISED: STANDARDIZATION DATA OF PERCENTAGE RETENTION AND COMPARISON OF RETENTION RATES WITH LOGICAL MEMORY SUBTESTS OF THE WECHSLER MEMORY SCALES - REVISED

A Dissertation

Submitted to the School of Graduate Studies and Research

In Partial Fulfillment of the

Requirements for the Degree

Doctor of Psychology

Stephen Fink

Indiana University of Pennsylvania

December 2008

Indiana University of Pennsylvania

The School of Graduate Studies and Research

Department of Psychology

We hereby approve the dissertation of

Stephen Paul Fink

Candidate for the degree of Doctor of Psychology

____________________ ____
David J. LaPorte, Ph.D.
Professor of Psychology, Advisor

____________________ ____
Dasun Luo, Ph.D.
Professor of Psychology

____________________ ____
Michael D. Franzen, Ph.D.
Associate Professor of Psychiatry
Drexel University College of Medicine

ACCEPTED

____ ____________________
Michele S. Schwietz, Ph.D.
Assistant Dean for Research
The School of Graduate Studies and Research

Title: Hopkins Verbal Learning Test - Revised: Standardization Data of Percentage Retention and Comparison of Retention Rates with Logical Memory Subtests of the Wechsler Memory Scales – Revised

Author: Stephen Fink

Dissertation Chair: Dr. David LaPorte

Dissertation Committee Member: Dr. Dasun Luo

Dr. Michael Franzen

The purpose of this study was to determine the utility of the Hopkins Verbal Learning Test-Revised (HVLT-R), as a brief neuropsychological screening measure of immediate and delayed verbal memory in a mild traumatic brain injury sample. From an archival database, neuropsychological test scores of 698 subjects with a recent mild-TBI were selected for the study.

Age and education effects were analyzed with nonparametric measures. Correlations between the neuropsychological tasks were conducted. Post-hoc exploratory analysis involved the creation of frontal lobe superior and impaired groups based on performance on two frontal lobe tasks: the Controlled Oral Word Association Test (COWAT) and Trails B from the Trail Making Test A and B. Independent sample t-tests were conducted to analyze the groups.

Results suggest that several narrow age groups published in the HVLT-R's test manual were not found with the current head-injured sample. Three age group categories emerged from the results: 18-29 years; 30-69 years; and 70+

years. Results of education level indicated no significant difference between subjects with a high school diploma or equivalency and subjects with less than a high school education. However, both groups significantly differed from subjects with more than a high school education.

The HVLT-R was compared with Logical Memory subtests from the Wechsler Memory Scales-Revised (WMS-R). Although correlations between the measure's various tasks were moderate, the HVLT-R identified a much higher number of subjects in the impaired range. Exploratory analysis involved creating groups based on the top and bottom quartiles of individuals who had taken the COWAT and Trails B, two frontal lobe tasks. The comparison of the HVLT-R and Logical Memory on the established "frontal lobe superior" and "frontal lobe impaired" groups revealed that both measures were able to identify subjects with impaired frontal lobe functioning and differentiate them from individuals with intact frontal lobes. The first trial of the HVLT-R was correlated with another test of focused attention and immediate memory, the Digit Span subtest of the WMS-R. Results suggested a moderate correlation between the tasks.

DEDICATION

I would like to dedicate my dissertation to the memory of my son, Stephen Ashton Fink, who unexpectedly passed away one day after I submitted my final copy of this dissertation to my committee. He was three weeks shy of his fourth birthday when he passed away. I had hoped to spend more quality time with him once “my paper,” as he called it, was complete, but fate can sometimes be very cruel. He has left an incredible void in our lives, and we will always miss his infectious laugh, his love for life, his extraordinary mind and his remarkable thirst for knowledge. He amazed us and brought us great joy every day of his brief yet memorable life. We will always love and miss our precious Ashton.

ACKNOWLEDGEMENTS

I want to first thank my Lord and Savior, Jesus Christ, for blessing me in so many ways and giving me hope that I could persevere even when I thought that hope might be lost.

I want to thank my beautiful wife, Alisha, for her constant love, patience, and encouragement throughout this process. Alisha, I am so blessed to have you in my life, and I love you very much.

I want to thank my son, Andrew, for involuntarily giving up so many hours and days with Daddy over the past year so that I could complete this dissertation. I know you are too young to comprehend the reasons for the missed time, but I promise I'll make it up to you. I love you.

I want to thank my parents, Bob and Phyllis Fink, and my in-laws, Chuck and Kathy Hutzler, for their support, encouragement, and countless hours of taking care of our sons so that I could concentrate on this project. I also want to thank my father-in-law, Fred Clark, for his concern, support, and wise counsel.

I want to thank our friends, Don Patchell, Heidi Dobish, and Jared and Rachel Detter, for their encouragement, support and willingness to proofread. I also want to thank my wife, mother, and mother-in-law for proof-reading my dissertation. Anyone willing to read large portions of someone else's dissertation deserves mucho gratitude.

I would like to thank my dissertation chair, Dr. David LaPorte for guiding me through this process as well as my committee members, Dr. Michael Franzen and Dr. Dasen Luo.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I: INTRODUCTION

Traumatic brain injury directly impacts millions of individuals in the United States and also affects countless others who care for or depend on individuals who suffer brain trauma. Others are affected by the high medical costs necessary to rehabilitate brain injury survivors as well as the cost to society in loss of occupation productivity because of the limitations many brain-injured individuals face which preclude their return to their former employment (Giles & Clark-Wilson, 1993).

Assessment of cognitive impairment in individuals who have suffered brain injuries is crucial in order for prognosis to be assessed and appropriate rehabilitative treatment to be implemented (Shutter, Jallo, & Narayan, 2000). Neuropsychologists conducting outpatient assessments spend hours with each patient thoroughly measuring cognitive domains such as orientation, attention, language, immediate and long-term memory, visuoperceptual functions, motor functioning, cognitive flexibility, and the ability to organize thoughts and plan appropriate behavior.

Unlike the outpatient environment, it is not generally feasible to administer a full neuropsychological battery in the inpatient setting. Pressure from healthcare provider organizations to decrease the time that patients receive treatment in the acute care setting makes lengthy test batteries impractical (Franzen, 1998). Not only is this true of the acute care settings, it also is an issue in rehabilitation settings.

An added difficulty in the rehabilitation setting is the demand placed on the patient’s schedule by treatment team members in several disciplines. Often in rehabilitation hospitals each treatment team member schedules time daily to work with each patient on their caseload. Neuropsychologists may find themselves limited to spending only brief periods of time daily with each patient in order to share the patient’s schedule with physical therapists, occupational therapists, speech pathologists, neurologists, and other clinicians involved in the patient’s care (Green, Stevens & Wolfe, 1997). The continuity necessary to complete a lengthy battery of neuropsychological tests is not practical in such situations.

Besides the difficulties with patients’ schedules, the neuropsychologist’s schedule often precludes working with one individual for several hours. Many neuropsychologists working in acute care settings and rehabilitation settings are responsible for screening several patients for cognitive difficulties in a single day, and the use of long, laborious psychological tests would not afford these neuropsychologists the opportunity to keep up with the high demand for such assessments.

A final drawback to full neuropsychological batteries in inpatient and rehabilitation settings is the impact long batteries have on the patient (Brandt & Benedict, 2001). Many patients are unable to tolerate an extensive neuropsychological battery requiring several hours to complete. Due to the recency of their head injury, many patients fatigue much too quickly for a time-consuming battery to be attempted. In such cases, the neuropsychologist may need to utilize a battery of tests consisting of brief screening tools used to measure several aspects of cognitive functioning within a circumscribed period of time. Further testing on an outpatient basis can be completed with a more complex neuropsychological battery once the patient is discharged from the hospital.

Due to the above limitations, neuropsychologists often administer a briefer neuropsychological screening while the patient is still in an acute inpatient setting in order to distinguish between patients who are in need of a more thorough evaluation and those whose cognitive functioning appears to be intact (Brandt & Benedict, 2001). Psychologists therefore are in need of brief and easily administered cognitive screening instruments that are psychometrically reliable and valid, and sensitive to the presence of cognitive impairment. Such screening measures allow the neuropsychologist to make recommendations for further assessment and treatment when appropriate. Assessment tools that have the ability to track cognitive change over time are also of great benefit in order for the neuropsychologist to assess patient progress during rehabilitative treatment.

Memory impairment is one of the most common sequelae of closed head injury (Bruce & Echemendia, 2003). It is no surprise, therefore, that one of the key areas that neuropsychologists attempt to measure subsequent to a brain injury is the assessment of short-term and long-term memory functioning. A multitude of assessment tools have been developed to pinpoint any deficits in a patient’s ability to encode and store information in memory and subsequently retrieve that information when needed. Some of these assessment tools require a great deal of time to administer.

The Hopkins Verbal Learning Test (HVLT), developed by Brandt (1991), was designed as a brief test of verbal learning and memory. The HVLT and its second edition, the HVLT-R (Brandt & Benedict, 1998; 2001) provide several advantages over similar verbal memory screening tools. It requires only a small amount of time to administer. Besides the brevity of administration, it is easily tolerated by patients, and it provides a screen for memory that considers the use of repetition in learning and memory. Some memory tests do not emphasize the importance of repetition in this capacity. Furthermore, unlike many other memory tests, the HVLT-R has six parallel forms. This is significant because of the frequent need to re-assess patients at specific intervals throughout the recovery process.

The Problem

Many studies have supported the utilization of the HVLT-R as a verbal memory screening test for dementia. Very little research, however, has investigated the clinical utility of this test with a head-injured population.

Purpose of the Study

The purpose of the current study was to investigate the clinical utility of the HVLT-R with subjects who experienced a recent mild closed head injury. The utility of the HVLT-R as a measure of immediate and delayed verbal recall for the mild head injured population was explored. Are differences in performance on the HVLT-R found with age and education level in a mild head injured population? How does the HVLT-R compare with another verbal memory task, the Logical Memory subtests of the Wechsler Memory Scales-Revised (WMS-R) with this population and, if differences exist between the verbal memory tasks, what might be the implications of these differences? Finally, could the first trial of the HVLT-R potentially be utilized as a measure of focused attention? To answer the above questions, archival data from an existing database was used and 698 subjects were included in the study.

CHAPTER II: REVIEW OF THE RELATED LITERATURE

Head Trauma

Closed head injuries and the brain trauma that often accompanies them have been a pervasive public health issue in the United States for decades. The Centers for Disease Control and Prevention (CDC, 2007) estimated that 1.4 million individuals annually in the United States alone sustain a traumatic brain injury. According to the CDC, approximately 50,000 of these individuals die as a result of their injuries. The government estimated that, of those who survive, 235,000 will be hospitalized and the remaining 1.1 million have their injuries treated by emergency department personnel and released. The CDC (2007) acknowledged, however, that the annual number of individuals who suffer a mild TBI but do not receive emergency medical care is unknown.

An estimated 5.3 million Americans currently suffer from the effects of a traumatic brain injury the extent of which results in long-term or lifelong need for assistance with activities of daily living (ADLs) (CDC, 2007). Of the 1.4 million annual head injuries in the United States, it is estimated that between 80% and 90% are classified as mild brain injuries (Hartlage, Durant-Wilson, & Patch, 2001; Shutter, et al. 2000). However, despite these individuals being labeled as having suffered a “mild” brain injury, many of these individuals will have a residual neurological disability (Shutter, et al., 2000). Therefore, proper neuropsychological assessment is a necessity for informing appropriate medical treatment and rehabilitative services.

In order to clear up any confusion that sometimes arises due to the habit of some neuropsychologists who use the terms “head injury” and “brain injury” interchangeably, it should be noted that there is a distinction. Goodyear and Umetsu (2002) pointed out that it is possible to damage the head without damaging the brain. It is also possible to suffer a brain insult without damage to the head. They recommended only using the term “brain injury” when there is physiological evidence of brain dysfunction and a documented alteration in mental status post–injury.

Classification of Head Injury

The importance of head injury classification must be emphasized because proper classification impacts the neuropsychological evaluation and treatment of individuals suffering a head trauma. Head injury can be classified in at least two diverse ways. First, head injuries are classified based on the post-trauma integrity of the skull. This method of classification distinguishes between open and closed head injuries. A second method of classification distinguishes between head injury severity. This second method distinguishes between mild, moderate, and severe injury. Each method of classification will be described below.

Open Versus Closed Head Injury

Head injuries can be classified as either closed head injuries or open head injuries. The term “open head injury” refers to an injury due to an impact to the skull causing the skull to fracture (Lucas, 1998). This often occurs due to the penetration of the skull by a missile-type object, such as a bullet. Because of the frequent association with a missile-type object, open head injuries are also called penetrating head injury. Once the integrity of the skull is breached, the brain is vulnerable to invasion of the brain tissue by any foreign object that has fractured the skull as well as fragments of the fractured skull.

As several authors have acknowledged (for example, Anderson, 1994; Gennarelli & Graham, 2005; Lezak, 1995), a closed head injury often involves injury to the brain due to a rapid acceleration-deceleration of the head. According to Lezak (1995), the acceleration/deceleration mechanism involves an injury in which the skull rapidly decelerates as the head comes into contact with a solid surface. Although the head is stopped, the brain continues to move forward, colliding into the hard, bony surface of the skull. This impact can and sometimes does occur at the sides or the back of the head, despite the typical forward motion of individuals traveling by vehicle or on foot. Consequently, the frontal and mesial-temporal lobes are located at high risk areas for brain injuries. It should be noted that the head does not necessarily have to come into contact with a hard surface for the brain to be at risk. In motor vehicle accidents, an individual's head rapidly decelerates from a very high rate of speed with no impact, yet the brain continues traveling at that high speed once the head has become essentially immobile. The soft tissue of the brain colliding with the hard surface of the skull is the result. This damage at the point of impact is called a coup lesion. After the brain collides with the skull and sustains contusions (bruising) from the impact, it bounces off the skull and recedes back until it collides with the hard, bony surface inside the skull opposite the point of impact; thus sustaining more contusions. The result of these contusions to the side of the brain opposite the point of impact is called contrecoup lesioning.

Hanlon, Demery, Martinovich, and Kelly (1999) noted that not all closed head injuries involve the acceleration/deceleration mechanism. In a study involving 100 subjects diagnosed with mild traumatic brain injuries, they differentiated between two types of acceleration/deceleration injuries: (1) subjects who experienced a rapid deceleration when their head struck an object (HSO) (n = 63); and (2) subjects who experienced a rapid deceleration but did not experience their head striking an object (HNSO) (n = 14). A third group in the study, however, included subjects who were stationary at the time of injury and experienced trauma due to the head being struck by an object (OSH) (n = 23). Analysis of variance and post hoc comparisons found significant differences between the three classification groups on several neuropsychological measures as well as vocational outcome. Injuries to the OSH group were found to result in significantly higher disruption to cognitive functioning and worse vocational outcome when compared with either acceleration/deceleration group.

Head Injury Severity

Head injuries are also classified as mild, moderate or severe. Methods to distinguish between these three levels of severity vary. Two prominent methods include assessment of alteration/loss of consciousness and assessment of posttraumatic amnesia.

Loss of consciousness. Loss of consciousness (LOC) has traditionally been the standard for classification of a brain injury (Bigler, 1990). Teasdale and Jennett (1974) developed the Glasgow Coma Scale (GCS) for the purpose of better assessing level of consciousness. The GCS was the first rating scale of level of consciousness to be empirically-derived and clinically validated (Bigler, 1990). The GCS is intended to be applied within 24 hours post-injury. Scores on the GCS range from 3 to 15 and are based on the assessment of three independent components: (1) eye-opening behavior; (2) best motor response; and (3) best verbal response (Lucas, 1998). As Table 1 indicates, the GCS classifies head injuries as mild, moderate or severe.

Table 1

Glasgow Coma Scale Classification of Head Injury Severity

Score	Severity
13 to 15	Mild
9 to 12	Moderate
8 or less	Severe

Adapted From: Teasdale and Jennett (1974)

Several authors and researchers have noted limitations in the ability of loss of consciousness, and the GCS in particular, to distinguish between individuals with and without head trauma. Lezak (1995) called attention to three limitations. She noted that when certain patients with little or no LOC are assessed with the GCS within 24 hours after trauma, misclassification may result. Patients without LOC may show no initial signs of mental status deterioration; however, a delay in deterioration may occur after 48 or more hours due to the effects of increased pressure on the brain from internal bleeding. A second limitation noted by Lezak is the discrepancy in LOC between right- and left-hemisphere injuries. She indicated that patients with penetrating wounds to the left-hemisphere of the brain are more prone to LOC than patients with injuries to the right hemisphere. She also noted that research has shown a discrepancy in the duration of coma between right- and left-hemisphere damaged patients. A third limitation expressed by Lezak is the impact alcohol intoxication has on GCS scores. There is a negative correlation between GCS scores and blood alcohol level (BAL) at the time of injury. Therefore, an intoxicated individual may have an artificially suppressed GCS score which subsequently improves several hours later as the individual becomes sober.

Lucas (1998) noted that, although the GCS is sensitive to moderate and severe head injuries and therefore is a useful predictor of neuropsychological outcome when injury falls in the moderate to severe range, it is not as sensitive to mild head injuries. The results of two recent studies (Iverson, Lovell, & Smith, 2000; Lovell, Iverson, Collins, McKeag, & Maroon, 1999) substantiate Lucas' concerns. Therefore, these two studies are further described below.

Lovell and his colleagues (1999) accessed an archival database from the Trauma Service at Allegheny General Hospital in Pittsburgh, Pennsylvania, to research the ability of loss of consciousness to predict performance on several neuropsychological tests in a large sample of head-injured individuals. The study included 383 subjects with mild head trauma, assessed with the GCS. A patient's GCS score upon admission to the trauma unit had to be between 14 and 15 in order to be included in the study. The researchers excluded patients who had skull fractures, posttraumatic amnesia, and/or abnormal CT-scans on the day of injury indicating swelling, bleeding or bruising of the brain. The researchers intended to relate their findings to LOC in sports and, therefore, they excluded any subjects older than 45-years. Subjects were separated into three groups: those who had experienced LOC post-injury, those who had not experienced LOC post-injury, and those who were uncertain as to whether or not they experienced LOC. The neuropsychological test battery consisted of the Galveston Orientation and Amnesia Test (GOAT), the HVLT, Trail Making Test (parts A and B), Controlled Oral Word Association Test (COWAT), the Wisconsin Card Sorting Test (WCST), and specific subtests from the Wechsler Memory Scale-Revised (WMS-R) which included the Logical Memory subtest, Visual Reproduction subtest, and Digit Span subtest. The researchers found no significant difference between the three groups on any of the neuropsychological measures. The surprising results did not support the use of GCS scores or, in general, LOC as accurate indicators of mild traumatic brain injury.

Iverson and colleagues (2000) conducted a study similar to the study of Lovell et al. (1999) on 195 patients and obtained similar results. Subjects were assessed through the Head Injury Trauma Service at Allegheny General Hospital. They were referred for a neuropsychological evaluation if any indication of head injury was evident, including LOC, retrograde or amnesia, or confusion post-accident. GCS scores were used to assess injury severity and only subjects with GCS scores between 13 and 15 were included. Excluded from the study were any subjects who had cerebral contusions or lacerations, or intracranial hematoma as evidenced on their day-of-injury CT-scan. The authors noted that subjects fulfilled the definition of mild-TBI as defined by the American Congress of Rehabilitation Medicine. The neuropsychological test battery consisted of the same tests as in the Lovell et al. (1999) study. Iverson et al. (2000) also divided subjects into the same three LOC groupings as Lovell et al. The results showed that the three groups did not significantly differ on any of the neuropsychological measures. The two studies cast serious doubt as to the utility of the GCS and LOC as indicators of mild TBI.

Posttraumatic amnesia. As an alternative to the GCS, some clinicians consider posttraumatic amnesia (PTA) to be a measurement of injury severity (Lezak, 1995). PTA occurs when an individual is amnestic to events immediately following a head trauma. As Table 2 shows, severity of head injury is classified based on the temporal length of PTA. Individuals who experience PTA for less than 1 hour subsequent to a head trauma are classified as mildly injured; PTA between 1 and 24 hours indicates moderate injury; PTA for longer than 24 hours is classified as a severe injury.

Table 2

Post-traumatic Amnesia Classification of Head Injury Severity

Hours of PTA	Severity
Less than 1 hour	Mild
Between 1 and 24 hours	Moderate
More than 24 hours	Severe

The Galveston Orientation and Amnesia Test (GOAT; Levin, O'Donnell, & Grossman, 1979) is a measure of mental status that primarily assesses orientation as well as anterograde and retrograde amnesia. The GOAT detects PTA by asking the individual to verbalize the first memory he or she recalls after the injury. GOAT scores range from 0 to 100 with higher scores indicating better functioning (Lucas, 1998). Scores between 76 and 100 points indicate normal functioning; between 66 and 75 indicate borderline functioning; below 65 indicates impaired functioning (Lucas, 1998).

Causes and Risk Factors of Head Injury

Several types of accidents have the potential of placing involved individuals at a high risk for a head injury. Recent statistics by the CDC (2008), however, indicate that falls are now the leading cause of TBI at 28% of all annual cases of TBI in the United States. Infants and children up to age 4-years-old and the elderly over the age of 75-years are at highest risk of experiencing a TBI from a fall. According to the CDC, MVAs now account for 20% of all head injuries annually in the United States with individuals between the ages of 15- and 19-years at highest risk. The CDC claimed that other causes of head injury include being struck by or against an object (19% of TBI cases annually in the United States), assaults (11% of TBI cases), and other or unknown cases (22% of cases).

Prevalence rates in several studies have revealed that individuals between the ages of 15 and 24 years are at the highest risk for suffering head injuries (Rimel, 1981; Sosin, Sniezek, & Thurman, 1996). Studies have also found two other age groups to be at a particular high risk: the elderly (over the age of 64 years) and children under the age of 10, particularly under the age of 5 years (Goldstein & Levin, 1990; Kraus & Chu, 2005).

Other factors that place individuals at a higher risk for head injury include previous head injury, alcohol intoxication, and the individual’s gender. Experiencing one head injury places the individual at a higher risk for subsequent head injuries. An individual is up to 3 times more likely to suffer a second head injury compared to the non-injured population (Anderson, 1994;

Kraus & Chu, 2005). Alcohol intoxication can impair driving ability, coordination, ambulation, and judgment; such factors increase the risk of experiencing one of the three major causes of head injury (motor vehicle accidents, falls, and/or assaults) (Kraus & Chu, 2005).Gender is also a factor because statistics have revealed that males are approximately twice as likely as females to suffer a closed head injury (Kraus & Chu, 2005).

Neuropathology

Lezak (1995) stated that trauma to the brain can result in both focal and diffuse injuries. The consequences of each type can be categorized as either primary or secondary. Bennett, Ditmar, and Ho (1997) described the distinction between primary and secondary injuries. Primary brain injuries are injuries that occur at the moment of impact and directly produce damage to the brain. In contrast, secondary brain injuries result from the physiological changes originally instigated by primary injuries. This can include hemorrhages, tissue swelling and an alteration in blood flow to the brain (Lezak, 1995). The primary and secondary effects of focal and diffuse axonal injuries will be described below.

Brain damage is often the result of what has been labeled “focal brain injury” (Bennett, et al. 1997). Focal injuries result from the collision between the brain and the skull. Furthermore, the inside of the skull is abrasive. When acceleration-deceleration and rotational forces are in action, damage may result from the soft brain tissue scraping against the bumpy surface of the skull. The frontal and temporal lobes are two brain regions especially susceptible to this type of injury (Lezak, 1995; Selzer, 1995). Bennett and colleagues (1997) agreed that the frontal lobes are vulnerable to this type of brain injury but asserted that the mesial-temporal lobe region is well insulated and, therefore, the hippocampus and other mesial-temporal brain structures are not at great risk.

Bennett and colleagues (1997) described another very serious category of brain injury as diffuse axonal injury (DAI). Whereas the brain is floating in cerebrospinal fluid and is only attached at the spinal cord; it is susceptible to rotational forces during head trauma. Hence, the brain does not simply bounce back and forth during an acceleration-deceleration injury; it also twists and rotates as it bounces back and forth. The brain damage resulting from this twisting effect can be significant. In contrast to the focal contusions at the points of impact in coup-contrecoup injuries, DAI leads to generalized brain damage. As the brain twists, axonal fibers are stretched and torn. The result is a number of microscopic lesions that may have an accumulative effect (Bigler, 1990; Gennarelli, 1986).

Gennarelli and Graham (2005) provided a thorough description of the potential secondary injuries that may result from a focal and diffuse brain injury. As they explain, the gradual onset of secondary injury effects can result in the deterioration of a patient who initially upon evaluation in the trauma center appears to only have mild injuries. As described above, brain contusions potentially occur at the instant of impact. Due to the contusions, brain tissue begins to swell and may expand to the point that swelling may cause damaging intracranial pressure. The brain is completely incased in the skull with the only outlet at the brainstem area where the spinal cord exits. If swelling occurs, there is no place for the brain tissue to expand and eventually the confinement by the skull causes the swelling to compress brain tissue with neuronal damage and death as a tragic result. If the swelling continues, the brain tissue may be forced through the small cranial opening near the brainstem, resulting in a herniation in the brainstem region that is potentially fatal.

Besides brain swelling as a potentially dangerous secondary effect, Gennarelli and Graham (2005) also noted intracranial hematomas (pools of clotted or partially clotted blood) are very common in brain injured patients who initially appear lucid, but deteriorate post-injury and constitute one of the most avoidable consequences of brain injury if detected early. As with the swelling of the brain described above, bleeding that flows and pools in an area of the brain is trapped within the skull and has no outlet. Therefore, the increase in bleeding results in building pressure and compression on brain tissue.

Between the skull and brain are three membranous coverings collectively referred to as the meningeal layers or meningeal coverings. The dura mater is closest to the skull, followed by the arachnoid, and the pia mater is closest to the brain. Gennarelli and Graham (2005) noted that hematomas generally are classified according to where they are anatomically located and can appear between the dura mater and skull (epidural hematoma), below the dura layer (subdural hematoma), below the arachnoid layer (subarachnoid hematoma), or in the brain tissue itself (intracerebral or intracerebellar hematomas).

The potential consequences of traumatic brain injury include damage to brain systems vital to several cognitive functions. As Bruce and Echemendia (2003) noted, one of the most common sequelae of traumatic brain injury is memory impairment. The next section provides a brief explanation of human memory models.

Human Memory

In order to evaluate the utility of neuropsychological assessment tools, it is necessary to first have a basic understanding of how human memory is structured. Ellis and Young (1988) defined memory as the capacity to register, retain, and retrieve information. Several models of human memory have been developed, including Waugh and Norman's (1965) two-component model consisting of "primary memory" and "secondary memory", Atkinson and Shiffrin's box model (1968), and Craik and Lockhart's (1972) theory of multiple levels of processing. A review of each of these models is beyond the scope and relevancy of this review. However, two theories, Atkinson and Shiffrin (1968) and Baddeley and Hitch (1974), have gained wide acceptance and, therefore, will be briefly described below.

The Atkinson and Shiffrin Model

Atkinson and Shiffrin (1968) proposed a basic model of memory functioning that consists of several components. Initially, information from the external environment is transformed at the body's sensory receptors; i.e., the sensory organs, into neural messages that are sent to the brain via an electro-chemical process in which the neural messages are transmitted from neuron to neuron. Sensory information is held for only a few seconds in what is known as sensory memory; it is subsequently discarded and forgotten or is sent to immediate memory as the individual attends to the message. In other words, in order for the message to pass from sensory memory to immediate memory, attention is required. When a stimulus is registered by the nervous system, the system then must form a representation of that stimulus. This representation in the nervous system is called encoding and requires the process of immediate memory for the representation to develop. The encoding of the information is vital to the later process of storage into long-term memory.

Atkinson and Shiffrin (1968) theorized that in short-term store the message is held for less than thirty seconds before it is forgotten unless the message is encoded for storage in long-term memory. Rehearsal of the information can result in keeping the message in short-term store beyond the thirty seconds and can also be useful during the encoding process into long-term storage. Once a message is encoded into long-term memory it may be stored permanently although memory decay is possible at this stage and may result in the message being forgotten. The final process in the memory system is retrieval of the message from storage in long-term memory back into short-term memory in order for the information to be used. The information may be recalled freely into short-term, memory. However, if an individual is unable to recall the information freely from long-term storage into short-term memory, the message may still be recognized and subsequently retrieved if choices are given to the individual rather than requiring that the individual recall without any retrieval cues.

Baddeley and Hitch's Working Memory Model

Baddeley and colleagues (Baddeley, 1986; Baddeley & Hitch, 1974) proposed a model of a working memory system which was more complex than Atkinson and Shiffrin's (1968) concept of short-term storage. Their model consists of three distinct subcomponents: the central executive, the visual sketchpad, and the phonological loop. The central executive is a central control structure responsible for processing and storage of information as well as the regulation of the flow of information in the working memory system. The visuospatial sketchpad is a domain-specific subsystem that allows for the temporary storage of visual and spatial images. The phonological loop is also a domain-specific subsystem responsible for the temporary storage and processing of verbal material, including material related to numbers, speech and words.

According to Baddeley (1994) and Gathercole (1997), there are two components to the phonological loop: a phonological short-term store and a subvocal rehearsal process. The phonological short-term store uses a phonological code that represents material. Because this code is subject to decay over time, the subvocal rehearsal process becomes important for the purpose of refreshing those decaying codes in the phonological store. Furthermore, the rehearsal process recodes inputs that are non-auditory into a phonological form so that the phonological store can hold these items. Gathercole emphasized, however, that this process of recoding the non-auditory inputs is not necessary for auditory speech because auditory speech is directly sent to the phonological store without the rehearsal process being enacted.

The concept of the phonological loop is an important aspect of the working memory model when considering word lists like the HVLT-R because the length of words on the list can be a determining factor in whether the words will be susceptible to decay or not. In essence, the longer the word, the longer time it takes to articulate the word subvocally and the greater the likelihood for decay of the phonological representations in the short-term store (Gathercole, 1997). This phenomenon was first investigated by Baddeley, Thomson, and Buchanan. (1975) and has been labeled the word length effect because the shorter the articulatory duration of the items in a word list the better the serial recall of the subject. Therefore, one-syllable and sometimes two-syllable words are easier to remember than words which contain several syllables. Gathercole (1997) explained that this phenomenon holds true, not only for word lists presented in the auditory modality, but also for word lists presented visually.

Baddeley et al. (1975) studied the word length effect to investigate whether the effect has to do with the number of syllables in the words or whether it has to do with the duration of time it takes to articulate the word. They compared word lists, each containing two-syllable words. The lists that contained words of short spoken durations were found to result in better recall than those with longer spoken durations. This finding has been debated, however, and it is still unclear whether the duration of the spoken word has a major effect on recall (Gathercole, 1997). It should be noted that Baddeley's working memory model is only one of several short-term memory models. However, it has been well-received and is an appropriate model to guide the present study.

Short-term Memory Components: Immediate and Working

Terminology can sometimes create confusion when several terms or phrases are used for the same psychological construct. An example is the difference between short-term memory, immediate memory, and working memory? Squire and Kandel (2008) reported that current theory in the field of cognitive psychology considers short-term memory to contain two major components: immediate memory and working memory. At the moment information is received, immediate memory is activated. However, immediate memory's storage capacity is considered to be limited to approximately seven items and is limited in time to approximately 30 second unless the information is rehearsed. In order for the information to last longer without memory decay or forgetting, the information must transfer to working memory. Transfer to working memory requires that the information be actively rehearsed. When this occurs, the rehearsed information can be retained in working memory for several minutes and potentially can be consolidated into long-term memory. Squire and Kandel's comparison fits with the well-established findings of Miller (1956) who found that seven items, plus or minus two, is the limited capacity available to immediate memory processing unless information is rehearsed and/or mnemonic devices are utilized to increase capacity as well as length of time.

Vallar and Papagno (2002) stated that neuropsychological tests designed to tap immediate memory generally utilize the serial span technique administered in either the visual or auditory modality. Examples of the auditory serial span techniques include the digit span techniques. Baddeley (2000) made the distinction between the concept of short-term memory as a stage in a memory model and the methods and techniques that have been developed to study and assess short-term memory. According to Baddeley, two techniques used to assess verbal short-term memory are memory span and free recall. Lezak (1995) recommended that every assessment of memory should include a task that measures immediate recall (Lezak termed this “span of immediate retention), as well as a measure of learning and efficiency of retrieval. For the purposes of this dissertation, we will use the term “short-term memory” to refer to the memory process as reflected in the stages of the memory model. The term “immediate memory” will be used to refer to the immediate recall of items from a neuropsychological memory task presented 30 seconds or less prior to recall in either the auditory or visual modality.

Attention and immediate memory deficits subsequent to traumatic brain injury are often associated with lesions to the prefrontal cortex (Lezak, 1995). The mesial-temporal region, however, has not been found to impact immediate memory. Many amnestic patients, most notably H.M. (Sidman, Stoddard, & Mohr, 1968), who have suffered lesions to the mesial-temporal lobes, are unable to store memories in long-term storage; however, the immediate memory abilities of these patients remains intact. Damasio and Tranel (Damasio, Eslinger, Damadio, Van Osesen, & Cornell, 1985; Tranel & Damasio, 2002) studied a patient who had suffered complete bilateral temporal lobe damage that prevented him from learning any factual knowledge, yet the patient's immediate memory was intact, as evidenced by his ability to retain information for up to 45 seconds.

In order to measure attention and immediate memory functioning, neuropsychologists utilize a variety of tasks. When verbal immediate memory and attention is the focus, tasks may include the initial trial(s) of a verbal word-list or paragraph-length story, or a test of digit span recall such as Wechsler's (1987, 1997). In the current study, Logical Memory-1 was compared with the first trial of the HVLT-R as well as the Total Recall score of the HVLT-R. The first trial of the HVLT-R was also correlated with the Digit Span subtest of the WMS-R.

Besides the effect that the prefrontal cortex has on immediate memory functioning, the prefrontal region is also highly associated with executive functioning. The following section explains the association between executive functioning and verbal immediate memory.

Relationship Between Executive Functioning and Verbal Immediate Recall

In the current study, the Hopkins Verbal Learning Test-Revised (HVLT-R) will be compared to other neuropsychological tests, including measures designed to tap what is known as "executive functions." Therefore, although a detailed examination is beyond the scope of the present study, a brief explanation of what constitutes executive functioning and the relationship between executive functions and verbal immediate memory, in particular, is warranted.

According to Goodwin (1989), executive functioning involves four components: the ability to formulate goals, make plans, carry out goal-directed plans, and complete the goals in an effective manner. The primary neuroanatomical region believed to be responsible for this executive system is the frontal lobes and, in particular, the prefrontal cortex (Rains, 2002). Therefore, since the executive system plays a supervisory role in a variety of cognitive domains, damage to the frontal lobes may result in impairment in attention, language, visuospatial functioning, personality functioning, and both visual and verbal memory (Anderson, 1994). Lezak (1995) noted that the impaired performance of patients with frontal lobe damage on verbal immediate recall tasks is at least partially due to the decreased ability of the frontal lobe patient to withstand interference, as well as decreased ability to order and organize what they learn and make use of contextual cues.

One of the most common concerns reported by patients who have suffered mild traumatic rain injuries is problems with memory (Bruce & Echemendia, 2003). Two of the most frequently used methods of measuring aspects of verbal learning and memory functioning are story recall tasks and verbal word-list tasks. The Logical Memory subtest of the Wechsler Memory Scales (WMS-R, 1987; WMS-III, 1997) is an example of a story recall task.

There are a number of verbal word-list tasks, including the Rey Auditory-Verbal Learning Test (RAVLT; Rey, 1964), the Buschke Selective Reminding Test (SRT; Buschke, 1973), the California Verbal Learning Test (CVLT; Delis, Kramer, Kaplan, & Ober, 1987), and the Hopkins Verbal Learning Test-Revised (HVLT-R; Brandt & Benedict, 2001). Despite certain similarities, each of these verbal list learning tests have unique characteristics including the number of words included in the list, the number of trials administered, and whether or not the words are able to be separated into semantic categories.

Tremont, Halpert, Javorksy, and Stern (2000) noted that despite the general belief in a relationship between executive functioning and memory, there is a paucity of studies on this relationship. They noted a subjective observation based on their clinical experience that patients suffering from executive dysfunction perform substantially better on story memory tasks than on verbal list learning tasks.

Tremont and colleagues (2000) conducted one of the first studies to investigate the association between executive functioning and verbal memory. They compared The California Verbal Learning Test (CVLT) with the Logical Memory subtest of the WMS-R. Archival data from a university-based medical center was obtained on 96 patients (48 male and 48 female) who had been referred for a neuropsychological evaluation. A battery of neuropsychological tests was administered to each subject including the CVLT and the Logical Memory. Also included in the battery were several measures of executive functioning. These included subjects’ scores on the Wisconsin Card Sorting Test

(WCST) perseverative responses, Trail Making Test Part B, the Similarities (SIM) subtest from the Wechsler Adult Intelligence Test-Revised (WAIS-R), and the Controlled Oral Word Association Test (COWAT). The executive functioning measures were used to create two groups for comparison. The Minimal Executive Dysfunction (MED) group (n = 44) scored in the impaired range on zero or one executive functioning measure. The Significant Executive Dysfunction (SED) group (n = 52) scored in the impaired range on two or three executive tasks. Patients who scored in the impaired range on four or five executive tasks were excluded from the study. Analysis included t-test comparisons on all of the memory tasks. Results revealed that the MED group performed significantly better on all of the scores on the CVLT, including Total of Trials 1-5, Short-Delayed Free Recall, Long-Delayed Free Recall, Short-Delayed Cued Recall and Long-Delayed Free Recall. There were, however, CVLT measures that did not show a significant difference. The MED and SED groups did not significantly differ on Percentage Retention, Semantic Cluster, Serial Cluster, Slope, Perseveration, Intrusion, Recognition and False Positive Errors.

In contrast to the CVLT results, Tremont and colleagues (2000) found that the MED and SED groups did not significantly differ on either Logical Memory I or II. Percentage Retention on Logical Memory also did not distinguish between the two groups. Tremont and colleagues concluded that their results support the argument that both types of verbal memory tests, verbal word lists and story recall, should be included in a comprehensive neuropsychological evaluation in order to allow the neuropsychologist to distinguish between direct memory impairment and memory impairment due indirectly to the effects of executive dysfunction.

Busch, Booth, McBride, Vanderploeg, Curtiss, and Duchnick (2005) conducted a study similar to the Tremont et al. (2000) study. They noted that Tremont and colleagues' conclusion that differences found in performance between the CVLT tasks and Logical Memory tasks based on executive dysfunction was premature because the SED group in the Tremont et al. study had more severe neurological disorders than the MED group. Therefore, the results may have been due to the severity of global cognitive impairment in the SED group rather than specific executive dysfunction.

To investigate this, Busch and colleagues (2005) administered a comprehensive neuropsychological battery to 193 subjects who were either active duty military personnel or military veterans. Only subjects who had sustained a non-penetrating head injury resulting in traumatic brain injury were included in the study. The battery included the Vocabulary, Similarities, Block Design, and Picture Arrangement subtests from the Wechsler Adult Intelligence Scale-Revised (WAIS-R); Logical Memory I and II, Visual Reproduction I and II, and Digit Span Backward from the Wechsler Memory Scale-Revised (WMS-R); California Verbal Learning Test (CVLT) Visual Spatial Learning Test (VSLT), Controlled Oral Word Association Test (COWAT), Trail Making Test – Parts A and B, Stroop Color-Word Test, Boston Naming Test (BNT), and the Wisconsin Card Sorting Test (WCST). Based on the number of executive functioning tests in which they scored in the impaired range, subjects were assigned to one of two groups. Scores from six executive tests were utilized to differentiate groups, including COWAT, Trail Making Test – Part B, Digit Span Backwards, Stroop Color-Word test, WCST Perseverative Responses, and WCST Set Failures. Similar to the study by Tremont and colleagues (2000) subjects who scored in the impaired range on zero or one measure of executive functioning were included in the Mild Executive Dysfunction (MED) group. Subjects who were impaired on two, three, or four measures of executive functioning were included in the Significant Executive Dysfunction (SED) group. Patients with impairment on five or more executive functioning measures were excluded from the study. Busch and colleagues (2005) also created matched subgroups based on the performance of subjects on other available neuropsychological tests measuring constructs other than executive functioning and memory. These measures included the Vocabulary and Block Design subtests of the WAIS-R, the BNT, and Trail Making Test – Part A.

Busch and colleagues (2005) analyzed the groups with a series of t-tests and found that the unmatched and matched groups differed in terms of performance on the neuropsychological tests. In the unmatched sample, the SED group performed significantly worse on every neuropsychological test. In contrast, the matched sample results found that the MED and SED groups differed significantly on each of the executive measures with the exception of Trail Making – Part B, but the two groups did not significantly differ on any of the other neuropsychological tests, including the memory tasks. A one year follow-up study was also conducted with 105 of the subjects to investigate whether the results during the acute phase of injury would resolve. Once again, the unmatched and matched samples differed. In the unmatched sample, the MED and SED groups significantly differed on every neuropsychological test except for Vocabulary and the BNT. In contrast, in the matched sample, the SED group scored significantly lower than the MED group on all executive functioning measures with the exception of Trail Making Test – Part B. The SED group also performed significantly worse on the visual memory tasks. However, there were no significant differences between the MED and SED groups on any of the other neuropsychological tests, including the verbal memory measures. Busch and colleagues suggested that their findings support the conclusion that, in the acute injury phase, differences between subjects on memory tasks are due to the overall severity of the neuropsychological impairment rather than severity of executive dysfunction. They further suggest that, after the acute phase ends (based on 1 year follow-up data) executive functioning does contribute to differences in visual memory performance, but performance differences on verbal memory tasks and other neuropsychological tasks are due to the severity of global cognitive dysfunction.

The two studies described above are important because, despite close similarities in procedures and design, the results and conclusions of Tremont et al. (2000) and Busch et al. (2005) were discordant. The current study investigated the association of another verbal list learning test, the Hopkins Verbal Learning Test-Revised, and the Logical Memory subtest. A brief description of the HVLT-R is provided below along with a review of the psychometric findings and clinical utility of the HVLT-R. For a thorough description of the HVLT-R administration procedures, see page 56 in Chapter 3 of this dissertation.

Hopkins Verbal Learning Test-Revised

Brandt (1991) developed the Hopkins Verbal Learning Test (HVLT) as a brief, repeatable screening measure of verbal learning and memory. It was designed for use in assessment situations in which lengthier and more comprehensive memory tools are impractical (Mitrushina, Boone, Razani, & D'Elia, 2005). The test consists of a list of twelve words semantically grouped into three categories. Benedict, Schretlen, Groninger, and Brandt (1998) introduced a revised version of the Hopkins Verbal Learning Test (HVLT-R). The HVLT-R kept the words lists from the original test and the procedures for administering the first three trials. The revised version is designed to assess learning efficiency, delayed verbal recall, delayed verbal recognition, and percentage retention. The revision solved one of the most evident limitations of the original HVLT by producing the delayed recall trial.

HVLT-R Normative Data

Several sources provide normative data for the HVLT-R. Brandt and Benedict (2001) developed the *Hopkins Verbal Learning Test-Revised: Professional Manual*. The normative sample consisted of 300 male and 879 female (n = 1,179) subjects ranging in age from 16 to 92 years. The mean age was 59-years-old. The reason for the high mean age (59-years in a sample spanning 16 to 92 years) is that age in the standardization group is highly skewed towards the older age groups, particularly the 70-79-year group in which roughly half of the total standardization sample (518 out of 1179 subjects) fall in one age group.

Years of education ranged from 2 years to 20 years with an average of 13.47 years of education. The manual provides norms, stratified by age, for Total Recall (total score summed from Trials 1, 2, and 3), Delayed Recall, Percentage Retention, and Recognition Discrimination Index. Data are available for eight age ranges beginning with 16-19-years olds, followed by age ranges stratified in 10-year increments, and ending with an 80+ range. Alternate norms stratified in 10-year increments beginning with 25-34 years and ending with 75-84 years are also available. The authors conducted a series of stepwise multiple regression analyses to investigate the influence of age, education, and gender on the four HVLT-R variables. They reported that age had the largest effect on each variable, followed by education level and then gender. However, despite gender and highest level of education completed making a statistically significant contribution to all four HVLT-R variables, the norms are not stratified by gender or level of education.

Benedict, et al. (1998) introduced normative data with the revised version of the HVLT. The standardization sample included 200 male and 341 female (n = 541) healthy subjects spanning ages from 17 to 88. Level of education ranged from 5 to 20 years with an average of 13.8 years. Norms were stratified by age range; however, unlike the 8 age ranges in Brandt and Benedict's (2001) professional manual, Benedict and associates (1998) found that 4 age ranges were distinct: 17-30 years, 31 to 54 years, 55-69 years, and 70-88 years. One advantage of the Benedict et al. (1998) normative study is that, not only did the authors provide norms for the Total Recall scores, Delayed Recall scores, Percentage Retention, and Recognition Discrimination Index, but also norms were provided for the initial three trials of the HVLT-R, a Learning score, and Recognition Measures such as True-Positives, False-Positives, and Response Bias.

Vanderploeg, Schinka, Jones, Small, Graves, and Mortimer (2000) also provided normative data specifically for an elderly population. The norms were based on a sample of 183 males and 211 females (n = 394) ages 60 to 85-years. The authors investigated the effects of age, education and gender in this elderly population via analysis of variance (ANOVA). The results of the ANOVA revealed significant effects for age and gender, but no effect for education or census status. The study resulted in normative data on only one age group age 60 to 84-years. Although the Vanderploeg et al. norms are restrictive in usefulness outside of the designated age range of 60 to 84 years, the norms have similar advantages as the Benedict et al.'s contribution (1998) because the authors provided norms for trials 1 through 3, Learning, Total Recall, Delayed Recall, Cued Recall, Percent retained, true positives, false positives and the Discrimination Index.

Norms for elderly African Americans were provided by Friedman, Schinka, Mortimer, and Graves (2002) based on a final data set of 108 male and 129 female (n = 237) elderly African American individuals, ages 60-84. The effects of age, gender, and education were examined. A multiple linear regression method revealed that age had a moderately large effect on subject performance of HVLT-R tasks. Subjects were then separated into two age groups and further analyses revealed that education and gender had significant moderate effects on performance on the HVLT-R tasks. A significant difference was found between individuals who had more than a high school education, those who had a high school education or equivalency, and individuals who had less than a high school education. The authors subsequently provided normative data for two age groups of elderly African Americans. The age groups were 60 to 71-years and 72 to 84-years. Furthermore, the authors provided gender and education normative corrections for the clinician to apply to each of the HVLT-R scores prior to looking at the age stratified norms.

In summary, each of the above sources of standardization data provides norms stratified by age. This is in contrast to the findings reported by Brandt (1991) when the original HVLT was introduced. He reported that there was no significant relationship between age and performance on the original HVLT and speculated that the lack of differentiation in age may have been due to the inclusion of only 7 individuals (out of n = 129) who were age 70 years or older. He also suggested that the normative group may have included a high number of healthy and well-functioning subjects since they were not a random sample.

Although all of the studies, with the exception of Brandt (1991), found that performance on the HVLT-R is affected by age, the studies are not in agreement as to how restrictive the range of age levels to be.

Several of the aforementioned normative studies also investigated the relationship between attained level of education and performance on the HVLT or HVLT-R. Brandt and Benedict (2001) found education-related effects; however, the normative tables in the professional manual are not stratified by level of education. Vanderploeg and colleagues (2000) examined the effects of level of education in their study and found no significant effect. In contrast, Friedman and colleagues (2002) did find a significant relationship between education and level of performance on the HVLT-R. They noted that Vanderploeg and his colleagues (2000) were studying a predominantly Caucasian sample. Friedman et al. (2002) were establishing norms for an elderly African American population. Benedict, et al. (1998) did not investigate level of education nor did they stratify their normative sample based on attained level of education. Therefore, there are disparate findings in the literature concerning whether level of education effects performance on the HVLT-R.

Reliability, Validity and Clinical Utility of the HVLT-R

Brandt (1991) studied the interform reliability of the six alternate forms of the original HVLT. He assessed 129 normal individuals ranging in age from 19- to 77-years who had achieved at least a high school education. One of the six alternate forms of the HVLT was randomly chosen for and administered to each subject. The Mini-Mental State Exam (MMSE; Folstein, Folstein, & McHugh, 1975) was also administered to 102 of the 129 subjects in order to screen for gross cognitive impairment. Subjects who completed the MMSE scored at least 25 out of 30 on the test which is indicative of intact performance. The results indicated that the six forms were highly intercorrelated.

In order to determine the stability of performance over time, Rasmusson, Bylsma, and Brandt (1995) studied the performance on alternate forms of the HVLT of 45 healthy elderly subjects ranging in age from 60- to 82-years who were each administered a neuropsychological battery on two separate occasions with a 9-month interval between assessments. As with Brandt's (1991) research on the interform reliability of the HVLT, the subjects in the Rasmusson and colleagues study were only included when their MMSE scores were greater than 25 at the time of both evaluations and when there was not more than a 2 point difference in scores between the first and second administration of the MMSE. The Logical Memory subtest of the WMS-R (Wechsler, 1987) was concurrently administered during the initial evaluation and again administered at the 9-month follow-up. The researchers found moderate stability for the recall and recognition measures on the alternate forms of the HVLT over the 9-month interval. They reported that the stability coefficients are compatible with test-retest reliabilities that have been reported for the California Verbal Learning Test (CVLT; Delis, Kramer, Kaplan, & Ober, 1986). Rasmusson and associates also found that the test-retest coefficient on the Logical Memory subtest of the WMS-R for their sample was similar to the test-retest coefficient reported in the WMS-R manual

(Wechsler, 1987) for people within the age range of their sample. The Logical Memory test-retest coefficient in the Rasmusson et al. study was comparable to the stability coefficient found for the HVLT recall measure in their research.

Shapiro, Benedict, Schretlen, and Brandt (1999) examined the construct and concurrent validity of the HVLT-R. The researchers used factor analysis and correlational analyses between the HVLT-R, other measures of memory, and general cognitive functioning. Participants were 55-years or older. Eighty-eight of the 302 subjects were normal adults and 214 were patients who had been diagnosed with a number of disorders, including 62 with probable Alzheimer’s disease, 38 with vascular dementia, 37 with mood disorder, 32 with schizophrenia, 11 with chronic alcoholism, 5 with Parkinson’s disease, and 29 subjects with a diagnosis of some other neuropsychiatric condition. In the analysis of the HVLT-R, the test was found to be isolated as a distinct factor loading on a different factor than other measures of neuropsychological tests, which include the Brief Visuospatial memory Test-Revised (BVMT-R; Benedict, 1997), a 30-item version of the Boston Naming Test (BNT; Kaplan, Goodglass, & Weintraub, 1983), the Trail Making Test (TMT; Reitan, 1958), the Visual-Motor Integration (VMI; Beery & Buktenica, 1982), word generation for the letters “S” and “P” (Borkowsi, Benton, & Spreen, 1967) and word generation for the categories “animals” and “supermarket items” (Rosen, 1980). Shapiro and associates (1999) also found a strong positive correlation between total recall on the HVLT-R and the Logical Memory-1 subtest of the WMS-R (.75), a strong correlation between the delayed recall trial of the HVLT-R and the Logical

Memory-2 subtest (.77) and a positive correlation of (.65) between percent retained on the HVLT-R and the Logical Memory percent retained. Weaker correlations were found between the total recall of the HVLT-R and the Visual Reproduction-1 (VR-1) subtest of the WMS-R (.54), the delayed recall of the HVLT-R and the Visual Reproduction-2 (VR-2) subtest (.69), and the percent retained on the HVLT-R and the Visual Reproduction percent retained. The results also indicated only a modest correlation between the HVLT-R and Verbal IQ (VIQ) or Performance IQ (PIQ) on the Wechsler Adult Intelligence Scale-Revised (WAIS-R; Wechsler, 1981). For total recall on the HVLT-R, correlations were .52 for VIQ and .49 for PIQ. For delayed recall on the HVLT-R, correlations were .36 for VIQ and .40 for PIQ. For percent retained on the HVLT-R, correlations were .14 for VIQ and .24 for PIQ.

Shapiro and colleagues (1999) also examined the extent to which the HVLT-R is able to discriminate between patients with dementia and healthy individuals. Two patient groups were included in the study. The first group consisted of 55 patients with probable Alzheimer’s disease (AD), and the second group consisted of 34 patients with vascular dementia (VaD). The first group ranged in age from 57-years to 88 years with a mean age of 76.1 years. Subjects in the second group ranged in age from 55-years to 85-years with an average age of 73.1 years. Average years of education between the two groups was relatively comparable with the AD group averaging 11.1 years of education (SD = 5.14, range 5-18) and the VaD averaging 11.4 years of education (SD = 3.7, range 4-20). Two groups of healthy individuals were chosen from a total sample of 445 subjects to act as a control groups for the clinical samples. The two normal groups were matched with the clinical groups on age and education level. The control group matched with the AD group consisted of 59 individuals with a mean age range between 70 and 88 years and an average age of 75.3 years. Their average level of education was 11.8 years (SD = 1.8, range 5-14). The control group that matched with the VaD group consisted of 37 individuals who ranged in age from 66 years to 88 years and with a mean age of 73.6 years. The mean level of education for this group was 11.4 years (SD = 2.3, range 5-14 years). Patients from each of the experimental groups completed a battery of tests including: the HVLT-R, Brief Visuospatial Memory Test-Revised (BVMT-R), the Trail Making Test, (TMT), Boston Naming Test (BNT), Visual-Motor Integration (VMI), Mini Mental State Examination (MMSE), word generation tests, and tests of emotional functioning. Subjects from the control groups completed less tests, but included the HVLT-R, BVMT-R, and TMT. The results showed that the normal subjects scored higher than the AD group on all measures of the HVLT-R. Similarly, the normal subjects from the second group scored significantly higher than the VaD group on the HVLT-R. The lone exception was the recognition task. The researchers also investigated the extent to which the HVLT-R discriminated between the AD and VaD groups. The results showed that the VaD group performed significantly better than the AD group on discrimination index and total recall. The results of this second study supported the discriminant validity of the HVLT-R as a measure of dementia in the elderly.

Studies have been conducted to compare the HVLT and the HVLT-R with other well-established assessment tools for memory and dementia. One such tool is the California Verbal Learning Test (CVLT).

Lacritz and Cullum (1998) studied the correlation between the HVLT and the CVLT in a group of 25 healthy elderly subjects (mean age = 70.7 years; SD = 9.3). Although non-significant correlations were found between the two measures on the first trial ($r = .30$, $p = .14$) and the second trial ($r = .31$, $p = .13$), significant correlations were found on the third trial ($r = .65$, $p = <.001$) of the HVLT and the fifth trial of the CVLT ($r = .65$, $p = <.001$), and total number of words learned ($r = .74$, $p = <.001$). The authors concluded that the relatively high correlation between the two tests on total words learned supports the use of the HVLT as a brief measure of learning. However, the authors noted concerns based on their results. By the third trial of the HVLT, the mean word recall was 10.16 out of a possible 12 words whereas the mean word recall on the CVLT by trial 5 was 11.90 out of a possible 16. The authors suggested that the additional items or trials on the CVLT made it a more sensitive measure of memory decline. However, the high number of subjects reaching the maximum score on trial 3 of the HVLT may not necessarily be abnormal in this sample of healthy elderly subjects because the sample consisted of very high functioning individuals as evidenced by their mean years of education (16.2 years, SD = 2.2). Given that the United States Census Bureau (2007) reported that 61% of individuals over the age of 64 years have 12 years of education or less, the sample in the Lacritz and Cullum (1998) may not accurately reflect the population on this particular dynamic. Lacritz and Cullum recognized the small sample size and high education level of their subjects limited the generalizability of their results and called for more comparative research between these two tests as well as other verbal memory tests using various clinical samples in order to determine the diagnostic utility of the HVLT.

They compared the performances of the HVLT-R and CVLT in a sample of 40 individuals diagnosed with probable Alzheimer's disease. Average age for the sample was 73.4 years (SD = 10.8) and the average level of education was 13.7 years (SD = 2.7). Subjects were assessed for level of dementia with the Dementia Rating Scale (Mattis, 1988) and were found to be within the mild to moderate range of impairment. A comprehensive neuropsychological battery that included both the CVLT and HVLT-R was administered to each subject. Pearson r correlations were analyzed. The results showed a modestly high and significant correlation between the total number of words learned on the HVLT-R and CVLT as well as the learning curve across measures. Retention rates were not significantly correlated. Short-delayed recall and 20-minute delayed recall were the highest correlated between the HVLT-R and CVLT of any other task on the tests. Recognition scores and false positive error scores were significantly correlated as well. However, the discriminability percentage was not statistically significant between the two measures. The authors acknowledged that the results that included modest and non-significant correlations may have been due to the restricted range of scores given the homogenous group studied. They stressed, for example, that most of the subjects were unable to recall a single word on the HVLT-R delayed recall trial, but most recalled at least one word from the CVLT. The authors theorized that this difference in performance may have been due to the extra 2 trials provided by the CVLT during the learning trial phase of administration. The authors concluded that the HVLT-R is a valid measure of verbal learning and memory. However, they recommended the CVLT for more complex diagnostic evaluations and the HVLT-R for brief neuropsychological assessments.

Research surrounding the clinical utility of the HVLT and the HVLT-R has focused primarily on forms of dementia. Frank and Byrne (2000) studied whether the HVLT could distinguish between normal older people and those with mild dementia. They assessed 26 subjects with a diagnosis of dementia based on criteria from the Diagnostic and Statistical Manual of Mental Disorders - Fourth Edition (DSM-IV; American Psychiatric Association, 1994). They also assessed 15 subjects with psychiatric diagnoses other than dementia and 15 normal controls. Types of dementia represented in the sample were 22 participants who had dementia of the Alzheimer's type, two participants who had vascular dementia, and two participants who had dementia due to multiple etiologies. According to the authors, all 26 subjects diagnosed with some form of dementia had only mild cognitive impairment based on their scores on the MMSE. Of the 15 participants with diagnoses of psychiatric disorders other than dementia, the authors reported that 10 subjects had major depression and five subjects had schizophrenia. The results of this study indicated that there were differences in abilities on the HVLT between the three groups of subjects. The mean HVLT

Total Recall score for the dementia subjects was 12.7 with a standard deviation of 3.9. In contrast, the mean Total Recall score for the psychiatric subjects was 20.4 with a standard deviation of 6.1, and the mean for the normal controls was 25.2 with a standard deviation of 2.8. The results of the study supported the research hypothesis that the HVLT could be used to distinguish between people experiencing symptoms of mild dementia and people who are not experiencing dementia.

Hogervorst, Combrinck, Lapuerta, Rue, Swales, and Budge (2002) investigated the sensitivity and the specificity of the HVLT to discriminate between patients diagnosed with dementia and controls. The researchers studied 114 control subjects and 82 subjects referred to their physicians by family members who suspected the subjects were suffering from dementia. All subjects were at least 55 years of age. The researchers excluded subjects with dementia if their score on the MMSE was below 9, and they excluded control subjects if their MMSE scores fell below 24. Of the 82 subjects with dementia, 68 were diagnosed with Alzheimer's disease, 6 with vascular dementia, 3 with Lewy body dementia, and 5 were diagnosed with other types of dementia. Results revealed that the Total Recall score on the HVLT discriminated well between the dementia subjects and the controls with a cut-off score of 14.5. The authors emphasized that they were chiefly concerned with specificity in their conclusion that a cut-off score of 14.5 was appropriate. They stressed that, if higher sensitivity is desired, a more appropriate cut-off score for the Total Recall on the HVLT would be 19.5.

Kuslansky et al. (2004) administered the HVLT and the Folstein Mini-Mental State Examination (MMSE) to 323 non-demented elderly individuals and 70 individuals who met the DSM-IV criteria for dementia. They found that the HVLT and MMSE only showed modest sensitivity to the detection of dementia. However, they noted that some of the elderly subjects in their control group had mild memory impairments, and others in the control group had mild cognitive deficits that may have indicated the early stages of dementia. They concluded that the HVLT scores should not be the only criteria for a dementia diagnosis and recommended further neuropsychological assessment to confirm the diagnosis.

A few studies have looked at the ability of the HVLT-R to accurately distinguish between mild closed head injury and non-head injured controls. Guskiewicz, Ross, and Marshall (2001) compared the ability of several neuropsychological measures as well as two measures of postural stability to accurately track recovery progress post concussion. Subjects were 36 collegiate athletes who sustained a concussion while participating in their sport and 36 recruited matched control subjects. Trail Making Test A and B, the Wechsler Digit Span subtest, Stroop Color Word Test, and the HVLT were used as the neurocognitive measures. They found significantly poorer performance on Trails B and Digit Span Backward in the head injured group, but the other neuropsychological measures, including the HVLT, were not significantly sensitive to the distinction between mild head injury and non-injured controls.

Iverson, et al. (2000) was described previously so the only addition to the review from the previous description will be the extent that the results pertain to the HVLT-R and brain injury assessment. For a more complete description of the logistics of this study, please see page 13. This study involved studied 195 mild head injured patients in the acute phase of recovery. Subjects were administered a neuropsychological battery that included the GOAT, TMT-A and B, HVLT, COWAT, WCST, Logical Memory, Visual Reproduction and Digit Span. Although the study was designed to investigate whether or not injury severity (as determined via neuropsychological test scores) could be differentiated based on whether subjects had loss of consciousness post-injury, another result showed that the majority of patients with mild head injury did not score unusually low on the test measures (defined as below the 10^{th} percentile) with the exception of the HVLT Total score. The study included three groups: a brain injured group positive for LOC, a brain injured group negative for LOC, and a brain injured group in which the status of an LOC was unknown undetermined. The percentage of subjects scoring in the impaired range on the HVLT-R was 52%, 41% and 54% respectively. The test closest to the HVLT in identifying individuals as impaired was the COWAT with 26.5%, 22.5% and 18%, followed by Trails B with 20%, 20% and 38% of the subjects identified as impaired. The researchers stated that this discrepancy may be due to educational differences. The HVLT normative sample had an education level with a mean of 13.8 years. The subjects in the Iverson, et al. sample had an education level with a mean of 12.1 years.

The clinical utility of the HVLT-R in discriminating between individuals with dementia and individuals without dementia has been established by several studies. However, there has been a paucity of research investigating the clinical utility of the HVLT-R with head-injured patients. The current study investigated the clinical utility of the HVLT-R as a screening measure of focused attention, immediate verbal and delayed verbal recall with subjects assessed for neuropsychological impairment subsequent to a recent mild closed head injury. The following section introduced the research hypotheses investigated in the current study.

Research Hypotheses

Hypothesis 1. It was predicted that lower scores on the Total Recall, Delayed Recall, and Percentage Retention tasks of the HVLT-R would be significantly associated with older age groups.

Hypothesis 2. It was predicted that lower scores on the Total Recall, Delayed Recall and Percentage Retention tasks of the HVLT-R would be significantly associated with a lower level of education.

Hypothesis 3. It was predicted that the Total Recall scores, Delayed Recall scores, and Percentage Retention scores on the HVLT-R would not be significantly correlated with the Logical Memory subtests of the WMS-R as a result of significantly poorer performance on the HVLT-R tasks than on the Logical Memory tasks.

Hypothesis 4. It was predicted that the first trial of the HVLT-R would be highly correlated with the Digit Span subtest of the WMS-R.

CHAPTER III: PROCEDURES

Participants

Participants were 698 patients who were assessed on an inpatient basis at Allegheny General Hospital in Pittsburgh, Pennsylvania, subsequent to experiencing a closed head injury. Due to various types of accidents, each subject was assessed for difficulties in orientation, attention, executive functioning, as well as short-term and long-term visual and verbal memory. Inclusion in the study required each subject to be diagnosed with a status post mild closed head injury. Exclusion criteria for the study included incomplete or non-existent scores on the HVLT-R, and a GOAT score of 65 or less which translates into an impaired performance (see page 56 for a more complete explanation of cognitive aspects assessed by the GOAT).

Table 3 displays the breakdown of subjects based on their assigned age group and sex. Subjects ranged in age from 18-years-old to 92-years-old, with an overall mean age of 42.2-years-old (SD = 18.49). Subjects were placed into one of eight groups based on their age at the time of the assessment. Males comprised 66.9% of the sample (N = 467) and females comprised 33.1% (N = 231).

Table 3

Number of Subjects Per Age Group and Sex

Age-group	N = 698	Males	Females
18-19	54	29	25
20-29	163	125	38
30-39	121	85	36
40-49	140	100	40
50-59	99	62	37
60-69	40	23	17
70-79	51	28	23
80+	30	15	15
Total	698	467	231

Table 4 displays the breakdown of subjects based on race in each age group. Overall, Caucasian subjects comprised 92.3% of the sample (N = 645), while African-American subjects comprised 7.6% (N = 53). There were no other racial groups represented.

Table 4

Number of Subjects Per Age Group and Race

Age-group	N = 698	Whites Subjects	Black Subjects
18-19	54	51	3
20-29	163	149	14
30-39	121	109	12
40-49	140	128	12
50-59	99	90	9
60-69	40	39	1
70-79	51	49	2
80+	30	30	0
Total	698	645	53

Table 5 shows the breakdown of subjects based on level of education in each age group. Subjects were categorized into one of three categories: (1) less than a high school graduate; (2) high school graduate; and (3) post-high school education. The post-high school category included subjects who had completed at least one year of college-level academic coursework. Therefore, subjects with only one year of college education were placed in the same category with subjects who had completed more than one year of college, as well as subjects who had an earned bachelors, masters, doctoral, or professional degree. Overall, 23.2% of the total sample size had less than a high school education (N = 162). High school graduates made up 43.1% of the sample (N = 301), while subjects with some level of education past high school made up 33.7% of the sample (N = 235).

Table 5

Number of Subjects Per Age Group and Level of Education

Age-group	N = 698	LHS[1]	HS[2]	PHS[3]
18-19	54	20	25	9
20-29	163	32	71	60
30-39	121	30	52	39
40-49	140	32	55	53
50-59	99	15	42	42
60-69	40	8	15	17
70-79	51	15	24	12
80+	30	10	17	3
Total	698	162	301	235

[1] LHS = Less than a High School Diploma

[2] HS = High School Diploma

[3] PHS = More than a High School Diploma

Level of education is perhaps misleading to some extent in the 18- to 19-year-old age group. Many 18-year-olds and even 19-year-olds may have less than a high school education because they are still working towards their high school diploma. Other subjects may have dropped out of school. No distinction was made between those subjects who are still in high school versus those students who have terminated their education prematurely. No distinction could be made between 18-year-old or even 19-year-old subjects who have not yet worked on post-high-school education and those who will not do so at any time. Therefore, for this one age group, level of education may be misleading.

Injuries and accidents that resulted in the head injuries of the subjects were tallied. The results can be found in Table 6. The majority of the head injuries resulted from motor vehicle accidents (63.6%). The next highest category of accidents involved falls of various types (22.5%). Victims of an assault via blunt force trauma made up the third-highest category of injury (4.9%). Other injuries grouped in Table 6 as “other” included pedestrian versus automobile accidents (2.9%), bicycle accidents (1.3%), bicycle versus automobile accidents (0.6%), injuries subsequent to gunshot wounds (0.7%), and other accidents (1.2%). There were also injuries sustained for unknown reasons (2.4%). Some of these unknowns were due to the psychometrists failing to record the circumstances surrounding the accident, while others were due to the subject being found unresponsive with no clear explanation for what happened to them.

Table 6

Frequency of Accidents/Injuries Among Subjects

Age	Auto	Fall	Assault	Other	Unknown
18-19 year olds (N = 54)	48	2	0	4	0
20-29 year olds (N = 163)	128	13	12	7	3
30-39 year olds (N = 121)	81	18	12	6	4
40-49 year olds (N = 140)	88	28	6	15	3
50-59 year olds (N = 99)	54	30	3	10	2
60-69 year olds (N = 40)	14	21	1	4	0
70-79 year olds (N = 51)	20	29	0	0	2
80 + year olds (N = 30)	11	16	0	0	3
TOTAL (N = 698)	444	157	34	46	17

Measures

Measurement tools utilized in this study consisted of a standard battery of neuropsychological tests administered to all patients suffering a closed head injury and subsequently treated at Allegheny General Hospital, Pittsburgh, Pennsylvania, on an inpatient status.

The full inpatient battery consisted of the following assessment measures: Hopkins Verbal Learning Test-Revised (HVLT-R); Logical Memory I and II subtests from the Wechsler Memory Scales-Revised (WMS-R); Visual Reproduction I and II subtests from the WMS-R; Digit Span subtest from the Wechsler Memory Scale-Revised (WMS-R); The Controlled Oral Word Association Test (COWAT); Animal Naming; Trail Making Test A and B; the Wisconsin Card Sorting Test (WCST); and a clinical interview. For the current study, the HVLT-R, WMS-R subtests, COWAT, and Trails B were utilized. (Full descriptions of these measures are provided below.)

Procedures

Subjects were admitted to the Department of Trauma Surgery at Allegheny General Hospital (AGH) on an inpatient basis subsequent to closed head injuries sustained via various injury modalities. Subjects were referred to the Department of Psychiatry at AGH by the Department of Trauma Surgery for neuropsychological screening. It is standard practice at the Trauma Services of AGH to refer all patients to the Department of Neuropsychology when there is any indication that the patient has suffered a head injury. Therefore, patients are referred for a neuropsychological evaluation if there is evidence that they suffered a loss or alteration of consciousness, experienced retrograde and/or anterograde amnesia, and/or exhibited other neurological signs such as disorientation or confusion. The Trauma Services staff evaluates the patient medically and cognitively, and if a head injury is suspected, the staff refers the patient to the Department of Psychiatry for an evaluation of several neuropsychological functions, including attention, immediate and delayed memory recall and recognition in the verbal and visual modalities, visual scanning and sequencing, and language and executive functioning.

After the referral is made, staff from the Department of Psychiatry trained to conduct neuropsychological testing performs a chart review. The evaluator then subjectively assesses whether the patient is conscious, alert, and able to complete neuropsychological testing with good effort. If so, the evaluator begins the assessment with a clinical interview and includes the Galveston Orientation and Amnesia Test (GOAT) in order to assess the patient’s orientation and both anterograde and retrograde amnesia. Items from the GOAT assess orientation to person, time, place and circumstance as well as questions concerning the last memory the subject has prior to the injury (assesses retrograde amnesia) as well as the first memories subsequent to the injury (assesses anterograde amnesia). The GOAT is based on a 100-point scale and a score of 65 and below is considered an impaired performance. Failure to recall events immediately prior to and subsequent to the injury results in a 20 point deduction. Therefore, to score below 66 points, an individual would not only have to display considerable difficulty with anterograde and/or retrograde amnesia, but would also have to answer several basic orientation questions incorrectly. When a patient displays great difficulty in successfully completing orientation tasks, the patient is likely to be too impaired to successfully perform on more complex cognitive tasks of attention, memory, and executive functioning. Therefore, if patients score in the impaired range on the GOAT, the evaluator will postpone testing until a passing score can be achieved. If the patient achieves a passing score on the GOAT, the evaluator proceeds with the neuropsychological battery that includes the Digit Span Forward and Backward subtests, Logical Memory and Visual Reproduction subtests of the Wechsler Memory Scales-Revised, the Hopkins Verbal Learning Test-Revised, the Controlled Oral Word Association Test, Animal Naming, Trail Making Test A and B, and the Wisconsin Card Sorting Test.

Subjects in the present study were administered the standard neuropsychological screening battery generally between 24- and 72-hours post-injury. On rare occasion when a subject was unable to be tested within the 24- to 72-hour period, subjects were tested no later than 7 days post-injury. The battery was administered and scored by doctoral-level psychology interns, doctoral-level psychology practicum students, and psychometrists employed and/or trained by AGH.

With IRB-approval from both Allegheny General Hospital and Indiana University of Pennsylvania, and in accordance with the Health Insurance Portability and Accountability Act of 1996 (HIPAA, Title II), archival records were obtained on subjects from the Department of Psychiatry at Allegheny General Hospital. No identifying characteristics were obtained. Records containing the scores of each subject were obtained from Neuropsychology Services at Allegheny Hospital. These records consisted of a face sheet constructed from the record containing the scores of the subject on the inpatient neuropsychological screening battery tasks, age, race, gender, education level, and handedness. Many of the records also indicated whether or not the subject had lost consciousness, and whether or not they were restrained or unrestrained (if the head injury occurred due to a motor vehicle accident). Also included was an explanation of the mechanism of injury (e.g., motor vehicle accident, fall, assault, etc.). Names and other identifying characteristics of the subjects were deleted. The results were compared with the normal control group provided by the *Hopkins Verbal Learning Test-Revised Professional Manual* (Brandt & Benedict, 2001) and with the normative data provided in the Wechsler Memory Scales-Revised Manual (Wechsler, 1987). The Professional Manual for the HVLT-R does not contain normative data for the first trial, so norms from the original HVLT (Brandt, 1991) were used. Data were analyzed using the SPSS Graduate Pack 13.0 for Windows statistical program.

Hopkins Verbal Learning Test-Revised (HVLT-R)

The HVLT-R (Brandt & Benedict, 2001) is an individually-administered verbal learning task for a word list presented in the auditory modality. It consists of a 12-noun word list, with four words in each of three semantic categories (e.g., gemstones, shelters, and animals). These words are presented verbally with no visual stimuli provided. The test consists of three initial trials involving the presentation of the word list by the administrator at an approximate rate of one word every two seconds.

After the examiner verbalizes the words on the first trial, the subject is asked to verbally recall as many of the words from the list as he or she can, in any order. Responses are recorded, and the entire list is then read again by the examiner (Trial 2) following the same format as on Trial 1. The subject is then asked to again recall as many of the words he or she can, in any order, from the word list, including the words that were already recalled after the administration of the first trial. The procedure is repeated in the same way for Trial 3. Once the responses are recorded for Trial 3, there is a 20-minute delay before the examiner continues with the delayed recall trial. After Trial 3, the subject is not informed that they will be asked later to recall the words without any cue.

For the delayed recall trial, the subject is reminded that there was a word list they learned previously and are then asked to recall any of the words that they can. This is the delayed recall trial (or free recall trial). Once this is completed, a delayed recognition trial (also called a forced-choice trial) is administered in which the examiner verbalizes a list of 24 words and the subject is asked to simply state “yes” or “no” after each word, indicating whether the subject believes that the word was on the initial word list or not. Each of the 12 original words is contained on the Delayed Recognition word list of 24. Also included are 6 words that, although not on the original list, are nevertheless related to one of the three semantic categories. The other 6 words are semantically-unrelated words.

The scores on the HVLT-R provide a measurement of a learning curve from Trial 1 to Trial 3, an initial Total Recall memory score, a Delayed Recall score, a Delayed Recognition score, and a calculation of False Positives including the semantically-related and semantically-unrelated words. The percentage of words the subject is able to retain from the initial trials to the delayed recall trial is also calculated. This is calculated by dividing the delayed recall raw score by the highest score achieved between either Trial 2 or Trial 3 of the initial trials.

Different scores reflect different aspects of memory function. The first trial recall score reflects immediate auditory memory, while the 20-minute delayed recall trial is a measure of recent (long-term) memory.

Logical Memory I and II (LM-I, LM-II)

The Logical Memory I and II subtests are part of the Wechsler Memory Scales-Revised (WMS-R) (Wechsler, 1987). The subtests require the patient to recall short paragraph-length stories from memory. These individually administered subtests, although listed as separate subtests, are interrelated as Logical Memory-II is the delayed recall trial of Logical Memory-I. Logical Memory-1 consists of two short paragraph-length stories (Stories A and B) presented in the auditory modality with no visual cues. The examiner reads Story A to the subject, at the conclusion of which the subject is expected to verbally recall as much as he or she can about the story. Story B is subsequently read to the subject, and the subject is again asked to verbally recall as much as he or she can about the story. After Logical Memory-1 is administered and responses are recorded, the examiner continues with the administration of other WMS-R subtests.

After a 30-minute delay, the examiner asks the subject if he or she recalls the stories that were previously read. The subject is asked, once again, to recall as much as he or she can about the two stories (this constitutes the delayed memory trial – Logical Memory-II). Responses to Logical Memory-II are recorded on the record sheet for later scoring. Logical Memory primarily assesses immediate verbal memory (Logical Memory-I) and delayed recall memory (Logical Memory-II).

Visual Reproduction (VR-I, VR-II)

As with the Logical Memory subtests, the Visual Reproduction I and II subtests of the WMS-R are actually interrelated despite being listed as separate subtests. During administration of Visual Reproduction-I, the examiner informs the subject that he or she will be presented with stimulus cards, one at a time, containing a drawing. Subjects are also informed that they will only have ten seconds to observe each card before it is taken away and they are asked to draw the figure. After the instructions are completed, the examiner presents the subject with the first card, a simple geometric figure, for a period of 10 seconds. The visual stimulus is then removed and the subject is prompted to draw the figure from memory. Once the subject has completed the first drawing, the examiner presents the second stimulus card and the process is repeated. The third and fourth stimulus cards are presented in the same manner as the initial cards. Although all four stimulus cards contain relatively simple geometric figures, the cards are presented in order of increased complexity.

After at least a 30-minute delay, the examiner asks the subject to draw the figures again, one at a time and in any order, from memory (this constitutes the delayed memory trial – Visual Reproduction-II). Visual Reproduction primarily assesses immediate visual memory (Visual Reproduction-I) and delayed recall visual memory (Visual Reproduction-II).

Digit Span (DSF, DSB)

The Digit Span subtest of the WMS-R is a measure of immediate auditory memory and focused attention. There are two aspects to this subtest, Digits Forward (DSF) and Digits Backward (DSB). The DSF task is administered first. Subjects are asked to repeat a series of numbers presented verbally by the examiner in the same order as presented. There are six items presented in a progressively more difficult pattern with an additional digit added on each trial level. There are two trials per each item and the task is not discontinued until the subject makes errors on both trials.

The DSB task is more complex than DSF because it requires the subject to hold the string of digits presented by the examiner in memory and then reverse the order when repeating the string back to the examiner. As with DSF, the DSB contains six items with two trials per item.

Digit Span Forward (DSF) primarily measures focused attention and immediate verbal memory. Digit Span Backward (DSB), a more complex task than DSF, is considered to be a measure of focused attention, divided attention, and immediate verbal working memory.

Trail Making Test A and B

Trail Making Test A (TMTA) requires the subject to draw a line connecting randomly arrayed numbered circles in sequential order. This task primarily taps visual scanning and sequencing and visuospatial functioning (Anderson, 1994). The more complex task, Trail Making Test B (TMTB) requires the subject to alternate between numbers and letters in sequential order, once again connecting the numbered/lettered circles by drawing a line. Although TMTB, like TMTA, taps visual scanning and visuospatial functioning, it additionally taps cognitive flexibility due to the shifting cognitive set, and, therefore, acts as a measure of executive functioning (Anderson, 1994).

Controlled Oral Word Association Test (COWAT)

The COWAT is an oral task that requires the subject to produce as many words beginning with a designated letter as he or she can within a given time period. This primarily taps word fluency and is an indicator of cognitive productivity, another executive functioning ability (Benton & Hamsher, 1989).

Data Analysis

To test the hypotheses that age groups and differences in education would be significantly associated with HVLT-R scores, Kolmogorov-Smirnov and Shapiro-Wilk tests of normality were conducted and scores on the HVLT-R were found to be significantly non-normal. Therefore, Kruskal-Wallis tests were used to analyze the association of age and level of education on each of the HVLT-R scores. Post hoc analysis consisted of a series of Mann-Whitney U tests to assess the differences between age groups and differences between levels of education.

To test the hypothesis that scores on the HVLT-R would not be significantly correlated with Logical Memory scores, Pearson r correlation coefficients were calculated. Frequency scores for impairment were then calculated for each HVLT-R and Logical Memory test. Pearson r correlations were also calculated between measures of frontal lobe functioning (COWAT and Trails B) and the memory tasks of the HVLT-R and Logical Memory. Fisher’s z transformations were calculated to determine the significance of differences found between correlation coefficients.

Exploratory analysis consisted of the creation of two groups based on subject’s scores on two frontal lobe tasks: COWAT and Trails B. Subjects who scored in the top quartile on both measures were included in one group and subjects who scored in the bottom quartile on both measures were included in the second group. A series of independent sample t-tests were conducted for each memory task. Effect sizes were also calculated. A second analysis consisted of creating two frontal lobe groups based only on scores on the COWAT. Inclusion in the first group required a score in the top quartile and in the second group required a score in the bottom quartile. Independent sample t-tests were then conducted on these two groups and effect sizes calculated.

To test the hypothesis that the first trial of the HVLT-R would be highly correlated with the WMS-R Digit Span subtest, Pearson r correlation coefficients were calculated. Frequency scores for impairments were calculated for each Digit Span subtest and the trial 1 of the HVLT-R.

To explore the ability of Logical Memory-1, Digit Span Forward and Digit Span Backward to predict intact versus impaired status on the Trial 1 and Total Recall Scores of the HVLT-R, forward stepwise logistic regression analyses were conducted.

CHAPTER IV: RESULTS

Determination of Normally Distributed Data

Due to each of the subjects in the sample suffering a closed head injury, it would be expected that they would represent a somewhat homogenous group in terms of memory functioning. Scores on the Total Recall, Delayed Recall, and Percentage Retention of the HVLT-R, therefore, would be expected to be somewhat skewed. Indeed, this was the case.

In order to test normality of the HVLT-R scores, Kolmogorov-Smirnov and Shapiro-Wilk tests of normality were conducted. Hopkins Total Recall, D (698) = 0.06, $p < .001$, Hopkins Delayed Recall, $D(698) = 0.10$, $p < .001$, and Hopkins Percentage Retention, $D(698) = 0.12$, $p < .001$ were found to be significantly non-normal. Therefore, nonparametric tests were conducted.

Of the 698 subjects in the study, 110 subjects (15.75%) scored zero on the Delayed Recall trial of the HVLT-R. In order to understand this low outcome for so many subjects, exploratory analyses were conducted. Not surprisingly, the greatest number came from the oldest age categories, even though subjects from each of the age groups contributed to the total number of zero scores. Subjects in the three oldest age categories (60-69, 70-79, and 80+ years) only make up 17.33% of the total sample, but accounted for 33.63% of the zero scores on the Delayed Recall trial. This may be partially due to the decreased performance of the elderly on memory tasks in general. The HVLT-R norms indicate a consistent decrease in cut-off scores between impaired and nonimpaired performance on each subsequent age group. In fact, a score of zero on the Delayed Recall trial, while still in the impaired range, is less than 3 standard deviations below the mean on the HVLT-R. This is the only age group in which a score of zero does not correspond to a T-score at or below 20. Table 7 provides details of the number of subjects and percentages of zero scorers for each age group.

Table 7

Frequency and Percentage of Subjects Scoring Zero on the Delayed Recall Trial of the HVLT-R

Age Group and Percent of Total Sample	Subjects Scoring Zero	Percentage of Total Zero Scores
18-19 year olds (n=54; 7.7%)	2	1.8
20-29 year olds (n=163; 23.4%)	19	17.3
30-39 year olds (n=121; 17.3%)	12	10.9
40-49 year olds (n=140; 20.1%)	22	20.0
50-59 year olds (n=99; 14.2%)	18	16.4
60-69 year olds (n=40; 5.7%)	9	8.2
70-79 year olds (n=51; 7.3%)	16	14.5
80+ year olds (n=30; 4.3%)	12	10.9
Total (n=698; 100%)	110	

Age and Education Differences on the HVLT-R

Given the non-normality of the HVLT-R scores, Kruskal-Wallis tests were employed to analyze the effect of age and level of education on HVLT-R scores. Table 8 displays the Kruskal-Wallis statistic for age and level of education on each of the HVLT-R scores. As Table 8 shows, Total Recall scores, Delayed Recall scores and Percentage Retention scores were each significantly associated with both age and level of education.

Table 8

Effects of Age and Level of Education on HVLT-R Scores

HVLT-R Score	K-W for Age	K-W for Level of Education
Total Recall	123.44**	35.83**
Delayed Recall	62.81**	25.79**
Percentage Retention	35.16**	16.38**

** Equals p value less than .001

Post hoc analyses consisted of a series of Mann-Whitney U tests in order to further explore where the differences lie in age group and education level. Appendix A provides results of these analyses for age group and Appendix B provides results for education level.

Hypothesis 1 predicted that differences in age groups would be significantly associated with the Total Recall scores, Delayed Recall scores, and

Percentage Retention of the HVLT-R. As Appendix A indicates, certain age groups did not show significant differences with other specific age groups on each of the HVLT-R tasks. Based on the particular non-significant findings between specific age groups, it would be appropriate to combine the eight age groups into three age groups: 18-29-years, 30-69-years, and 70+years. This finding is consistent with several other neuropsychological tests that have generated rather wide age ranges in the normative samples, an example of which would be the normative data for the COWAT (Tombaugh, Kozak & Rees, 1996; as cited in Spreen & Strauss, 1998).

Hypothesis 2 predicted that differences in education level would be significantly associated with the Total Recall scores, Delayed Recall scores, and Percentage Retention of the HVLT-R. As Appendix B shows, individuals with more than a high school education differed significantly from both the less than high school education group and the high school education group. However, those with less than a high school education and those with a high school education did not significantly differ. Therefore, it would be appropriate to combine these two groups. This would result in two groups based on education level: group 1 would have 12 years of education or less and group 2 would have 13 years of education or more.

Relationship Between the HVLT-R and the WMS-R Subtests

Hypothesis 3 predicted that the Total Recall scores, Delayed Recall scores, and Percentage Retention scores on the HVLT-R would not be significantly correlated with the Logical Memory subtests of the WMS-R. Means and standard deviations for the measures of the HVLT-R and the Logical Memory and Visual Reproduction subtests of the WMS-R can be found in Table 9.

Table 9

Means and Standard Deviations for HVLT-R and Specific WMS-R Subtests

Measure	Mean	SD
HVLT-R Scale 1	4.99	1.61
HVLT-R Total Recall Score	19.32	5.56
HVLT-R Delayed Recall Score	5.11	3.36
HVLT-R Percentage Retention	60.31	35.12
Logical Memory – I	18.09	7.16
Logical Memory – II	12.73	7.47
Logical Memory Percentage Retention	66.54	27.01
Visual Reproduction-I	29.27	8.50
Visual Reproduction-II	21.89	11.06
Visual Reproduction Percentage Retention	70.48	26.85

Pearson r correlation coefficients between the HVLT-R scores, the Logical Memory subtests and the percentage retention on the Logical Memory and Visual Reproduction of the WMS-R are located in Table 10. Each of the correlations were statistically significant ($p = 0.01$). As expected, the two verbal memory tasks involving percentage retention, Logical Memory (LM-PR) and HVLT-R (H-PR), were more highly correlated ($r = .519$) than either verbal memory task correlated with the Visual Reproduction percentage retention category (VR-PR). The correlation between VR-PR and its fellow WMS-R task (LM-PR) was higher at $r = .400$, than it was with the H-PR task ($r = .323$).

Table 10

Intercorrelations for HVLT-R Scores and WMS-R Subtest Scores

Measure	LM-1	LM-2	LM-PR	VR-PR	H-1	H-TR	H-DR	H-PR
LM-1	___	.842**	.362**	.406**	.565**	.638**	.568**	.460**
LM-2		___	.733**	.489**	.520**	.634**	.658**	.566**
LM-PR			___	.400**	.301**	.417**	.517**	.519**
VR-PR				___	.360*	.489**	.444**	.323**
H-1					___	.845**	.538**	.372**
H-TR						___	.709**	.484**
H-DR							___	.890**
H-PR								___

** Correlation is significant at the 0.01 level (2-tailed)

Note: LM-1 = Logical Memory – 1; LM-2 = Logical Memory – 2; LM-PR = Logical Memory Percentage Retention; VR-PR = Visual Reproduction Percentage Retention; H-1 = HVLT-R Trial 1; H-TR = HVLT-R Total Recall; H-DR = HVLT-R Delayed Recall; H-PR = HVLT-R Percentage Retention

Immediate Memory

Immediate memory refers to a complex process which involves a variety of steps and is limited in both capacity and duration. For an explanation of this process, refer to page 22 through 25 of the current study.

As Table 8 indicates, the initial Logical Memory subtest (LM-1) was significantly correlated with the first trial of the HVLT-R (H-1) (r = .565). The LM-1 subtest correlated higher with the Total Recall of the HVLT-R (H-TR) (r = .638), even though the H-TR benefits from a learning curve after three initial trials of the HVLT-R word list. LM-1, in contrast, does not repeat the same story and, therefore, it generally would not be considered to benefit from a learning curve. The reason for the higher correlation is because a score that represents the accumulation of scores from several trials is more reliable than any individual score (sub-score) that is a part of the total score. Because the total score is more reliable, it will have a higher correlation with another measure even if the total score does not reflect the construct to the degree that an individual sub-score does. The first trial of the HVLT-R (H-1), for example, may be considered the comparable immediate memory task to the Logical Memory-1. Although immediate verbal memory is required to perform well on the Total Recall of the HVLT-R, H-TR is considered to be a measure of learning efficiency.

Exploratory analysis was conducted in order to determine the frequency of each immediate verbal memory task's categorization of subjects from the acute mild head-injured sample as either experiencing memory difficulty to the point of placement in the impaired range or not. The raw scores for the Logical Memory-1 subtest were compared with norms from the test manual (Wechsler, 1987), and percentiles for the subjects were obtained. Likewise, the raw scores from the Total Recall of the HVLT-R were compared with normative data found in the test's manual (Brandt & Benedict, 2001), and T-scores for the subjects were obtained. A performance at or below the 2nd percentile (corresponding to a T-score of 30 or a z-score of -2.0) on psychological tests is considered to be very poor performance (Spreen & Strauss, 1998) and alerts the psychologist to potential impairment in the cognitive ability being assessed. Performance on immediate memory tasks is held to this standard as well.

Trial 1 of the HVLT-R (H-1) did not have norms provided in the professional manual; however, normative data based on the original HVLT, published separately by the test author (Brandt, 1991) were obtained for the purpose of comparing the H-1 scores of the current study's subjects with a normative group. The norms for H-1 were not separated based on age, gender, race, or level of education. Instead, all subjects are compared with a table that considers five or more correct responses on H-1 to be within the normal range, four correct responses to be in the mildly impaired range, between two and three correct responses to be within the moderately impaired range, and one or less correct responses to be in the severely impaired range.

All 698 subjects in the present study had complete sets of HVLT-R scores. However, 55 of the subjects were not administered the Logical Memory-1 subtest, and therefore their scores were not included in the comparison.

Once T-scores or percentiles were obtained for each subject, frequency of scoring in the impaired range was compared for each age group between the H-TR versus LM-1 versus H-1. For the purposes of the comparisons, impairment (also considered to be the defective range) falls at or below two standard deviations below the mean (corresponding to a T-score below 30 or a percentile rank below the 3rd percentile) as is the standard in neuropsychological assessment (Spreen & Strauss, 1998).

The results of the comparison between H-1, H-TR, and LM-1 can be found in Table 11. As these results indicate, subjects in the current study with a diagnosis of status post mild closed head injury were over five times more likely to be assessed as having memory impairment when administered the HVLT-R Total Recall than the Logical Memory -1 subtest of the WMS-R. Furthermore, just by administering the Trial 1 of the HVLT-R, subjects were almost four times more likely to be assessed as having memory impairment than they would be if only the Logical Memory-1 subtest were administered.

Table 11

Frequency of Impairment Between H-1, H-TR, and LM-1 (N = 643)

Age Group	H-1	H-TR	LM-1
18-19 years old (N = 50)	8	30	9
20-29 years old (N = 152)	39	67	18
30-39 years old (N = 112)	30	41	9
40-49 years old (N = 127)	33	59	4
50-59 years old (N = 97)	34	52	7
60-69 years old (N = 37)	18	19	3
70-79 years old (N = 46)	32	20	3
80+ years old (N = 22)	19	12	4
TOTAL	213	300	57

Overall, H-TR identified 300 subjects (46.7% of the total subjects compared) as falling in the impaired range. The H-1 identified 213 subjects (33.1%) as having impairment. In contrast, the LM-1 subtest only identified 57 subjects (8.9%) as experiencing memory impairment.

As Table 12 shows, the Total Recall score of the HVLT-R identified a higher percentage of subjects in each age group as experiencing memory impairment when compared with the performance of the LM-1. Likewise, Trial 1 of the HVLT-R identified a higher percentage of subjects in each age group as experiencing memory impairment than the LM-1, with the lone exception being the 18-19 year age group in which the LM-1 placed more subjects than the first trial of the HVLT-R in the impaired range. As these results indicate, assessing a head-injured patient within a week post-injury as having memory difficulty occurs with greater frequency on the initial HVLT-R tasks than on the Logical Memory-1 subtest of the WMS-R.

Table 12

Percentage of Subjects Per Age Group Identified as Impaired Between H-1, H-TR, and LM-1 (N = 643)

Age Group	H-1	H-TR	LM-1
18-19 years old (N = 50)	16.0	60.0	18.0
20-29 years old (N = 152)	25.7	44.8	11.8
30-39 years old (N = 112)	26.8	36.6	08.0
40-49 years old (N = 127)	26.0	46.5	03.1
50-59 years old (N = 97)	35.1	53.6	07.2
60-69 years old (N = 37)	48.7	51.4	08.1
70-79 years old (N = 46)	69.6	43.5	06.5
80+ years old (N = 22)	86.4	54.6	06.5
TOTAL	33.1	46.7	08.9

HVLT-R Trial 1 and Digit Span

Hypothesis 4. It was predicted that the first trial of the HVLT-R would be highly correlated with the Digit Span subtest of the WMS-R. The Digit Span subtest was included in the WMS-R and WAIS-R as a measure of focused attention and immediate memory. The subtest is divided into a Digits Forward task and a Digits Backward task. A subject's score on Digits Forward and Digits Backward are then combined to obtain an overall Digit Span score. Digits Forward is a measure of focused attention. In contrast, Digits Backward taps working memory.

In order to investigate the potential for the first trial of the HVLT-R to tap functions similar to Digit Span, HVLT-R trial 1 was compared with the Digit Span subtest of the WMS-R. Means and standard deviations can be found in Table 13.

Table 13

Means and Standard Deviations for HVLT-R Scores and Digit Span Scores

Measure	Mean	SD
HVLT-R Trial 1	4.99	1.61
Digit Span	12.81	3.77
Digit Span Forward	7.55	2.11
Digit Span Backward	5.24	2.13

Pearson r correlation coefficients between the first trial of the HVLT-R and the Digit Span scores are located in Table 14. The correlations between the first trial of the HVLT-R and each of the Digit Span categories (Forward, Backward, and Total) were statistically significant (p = 0.01). Although only a moderately high positive correlation, the first trial of the HVLT-R correlated highest with the total raw score of the Digit Span subtest (r = .570). The next highest correlation between the first trial of the HVLT-R and the Digit Span categories was with Digits Forward (r = .562). The correlation between the HVLT-R first trial and the Digits Backward was the lowest of the Digit Span categories (r = .453).

Table 14

Intercorrelations for HVLT-R and Digit Span Subtest

	HVLT-R	Digit Span		
Measure	Trial 1	Forward	Backward	Total
HVLT-R Trial 1	___	.562**	.453**	.570**
Digit Span Forward		___	.577	.886**
Digit Span Backward			___	.888**
Digit Span Total				___

** Correlation is significant at the 0.01 level (2-tailed).

Exploratory analysis was conducted to compare the frequency of assessing subjects as impaired between the first trial of the HVLT-R and the Digits Forward and Digits Backward subtests. The frequency and percentages of impairment on the three tasks is found in Table 15. Although the magnitude is not as great, Table 15 displays similar results found in Table 10. Overall, 35% of the head injured subjects scored in the impaired range on the first trial of the HVLT-R. In contrast, results of the Digit Span Forward and Backward scores shows that 13.2% and 15.1% of clients, respectively, scored in the impaired range. The results of the Digit Span Total Score shows that 5.6% of the subjects scored in the impaired range. It should be noted that the test manual for the WMS-R provides normative data for Digit Span Forward and Digit Span Backward. However, the WMS-R manual does not provide norms for Digit Span Total Score. The normative data provided in the Wechsler Adult Intelligence Scale-Revised (WAIS-R; Wechsler, 1981) does provide scaled scores for Digit Span Total Score. The Digit Span subtest on the WMS-R and WAIS-R are structured the same; however the digits utilized are not exactly the same. The results in Table 15 and 16 that refer to Digit Span Total Score are based on the scaled score norms found in the WAIS-R manual.

Table 15

Frequency (and Percentages) of Impairment on H-1, Digits Forward, Digits Backward, and Digits Total Score

Age Group	Hopkins-1		Digits Forward		Digits Backward		Digits Total	
18-19 years old (N = 54)	9	(16.7)	5	(9.3)	5	(9.3)	1	(1.9)
20-29 years old (N = 163)	45	(27.6)	17	(10.4)	15	(9.2)	7	(4.3)
30-39 years old (N = 120)	34	(28.3)	19	(15.8)	18	(15.0)	6	(5.0)
40-49 years old (N = 139)	40	(28.8)	27	(19.4)	28	(20.1)	7	(5.0)
50-59 years old (N = 99)	35	(35.4)	13	(13.1)	18	(18.2)	6	(6.1)
60-69 years old (N = 40)	19	(47.5)	8	(20.0)	4	(10.0)	4	(10.0)
70-79 years old (N = 51)	37	(72.5)	1	(2.0)	11	(21.6)	3	(5.9)
80+ years old (N = 30)	25	(83.3)	2	(6.7)	4	13.3)	3	(10.0)
TOTAL (N = 696)	244	(35.1)	92	(13.2)	105	(15.1)	39	(5.6)

Table 16

Percentage of Impairment on H-1, DSF, DSB and DST With and Without 60+ year old Subjects

Test	18 to 80+ year olds (N = 696)	18 to 59-year-olds (N = 575)
HVLT-R Trail 1	35.1	28.4
Digit Span Forward	13.2	14.1
Digits Backward	15.1	15.0
Digits Total	5.6	5.0

Logistic Regression Analyses

Separate forward logistic regression analyses were conducted with the Trial 1 of the HVLT-R and the Total Recall Score of the HVLT-R as dependent variables. The analyses were performed in order to determine which independent variables (Logical Memory-1, Digit Span Forward, Digit Span Backward, and interaction effects between the variables) are predictors of status (intact or impaired) on the dependent variables. In the first analysis, the First Trial of the HVLT-R was entered as the dependent variable. Results from this logistic analysis are shown in Table 17. Regression results indicated the overall model fit of two predictors (interaction between the three predictors and interaction between Digits Forward and Logical Memory-1) was statistically significant in distinguishing between intact and impaired performance on Trial 1 of the HVLT-R (χ^2 (2) = 81.231, $p < .001$) and accounted for a substantial amount of variance (Cox & Snell R-Square = .119; Nagelkerke R-square = .166). Regression coefficients are presented in Table 17. *Wald* statistics indicated that the interaction between Digits Forward and Logical Memory-1 and the interaction between Digits Forward, Digits Backward, and Logical Memory-1 significantly predicted intact versus impaired performance on Trial 1 of the HVLT-R. However, the odds ratios for these predictor variables were very small.

Table 17

Regression Coefficients for Predictor Variables and HVLT-R Trial 1

	B	*Wald*	*df*	*p*	Exp(B)
DB x DF x LM-1	-.874	6.497	1	< .05	.417
DF x LM-1	-1.045	12.247	1	< .001	.352
Constant	.720	12.894	1	< .001	2.054

Table 18 provides a contingency table of the model's predictive value. The model predicted that 113 (18%) of the subjects would be impaired on Trial 1 of the HVLT-R. Of these, 76 scored in the impaired range giving the model a positive predictive value of 67%. A total of 527 subjects were predicted to score in the intact range and this proved to be the case in 393, providing a negative predictive value of 75%. The overall correct prediction rate of the model was 73%. A total of 430 subjects scored in the intact range and the model correctly predicted 393 of these, giving a sensitivity of 91%. Two-hundred ten subjects scored in the impaired range and the model correctly predicted 76 of these, giving a specificity of 36%.

Table 18

Contingency Table of Predicted and Actual Impairment Outcomes on HVLT-R Trial 1

	Actual Impaired Score	Actual Intact Score	Total
Predicted Impaired Score	76	37	113
Predicted Intact Score	134	393	527
Total	210	430	640

In the second forward logistic regression analysis, the Total Recall Score of the HVLT-R was included as the dependent variable. Regression coefficients for this analysis are presented in Table 19. The overall model fit of two predictors (interaction Digits Backward and Logical Memory-1 and interaction between Digits Forward and Logical Memory-1) was statistically significant in distinguishing between intact and impaired performance on Trial 1 of the HVLT-R (χ^2 (2) = 119.095, p < .001) and accounted for substantial amount of variance (Cox & Snell R-Square = .170; Nagelkerke R-square = .226). *Wald* statistics indicated that the interaction between Digits Forward and Logical Memory-1 and the interaction between Digits Backward and Logical Memory-1 significantly predicted intact versus impaired performance on Trial 1 of the HVLT-R. However, very small odds ratios were obtained for the predictor variables.

Table 19

Regression Coefficients for Predictor Variables and HVLT-R Total Recall Score

	B	*Wald*	*df*	*p*	Exp(B)
DB x LM-1	-1.249	18.968	1	< .001	.287
DF x LM-1	-1.991	29.650	1	< .001	.136
Constant	2.840	59.183	1	< .001	17.118

Table 20 provides a contingency table of the model's predictive value. The model predicted that 165 (26%) of the subjects would score in the impaired range on the Total Recall score of the HVLT-R. Of these, 139 scored in the impaired range resulting in a positive predictive value of 84% for the model. A total of 475 subjects were predicted to score in the intact range and this was the case in 284, providing a negative predictive value of 60%. The overall correct prediction rate of the model was 66%. A total of 310 subjects scored in the intact range and the model correctly predicted 284 of these, giving a sensitivity of 92%. Three hundred thirty subjects scored in the impaired range and the model correctly predicted 139 of these, giving a specificity of 42%.

Table 20

Contingency Table of Predicted and Actual Impairment Outcome on Total Recall Score

	Actual Impaired Score	Actual Intact Score	Total
Predicted Impaired Score	139	26	165
Predicted Intact Score	191	284	475
Total	330	310	640

The results of the two logistic regression analyses indicate that there is significant overlap between the impairment on the HVLT-R measures and the predictor memory tasks. However, there is a considerable portion of impairment on Trial 1 and Total Recall of the HVLT-R that is not accounted for by the other memory measures. This may indicate that the HVLT-R Trial 1 and Total Recall scores are measuring a unique aspect of immediate verbal recall and learning when compared with the Digit Span tasks and Logical Memory-1.

Long-Term Memory

In contrast to immediate memory, long-term memory is considered to have an unlimited capacity, and information stored in long-term memory potentially has an unlimited duration. This ability to store information is dependent on the functioning of the hippocampus as well as other structures in the mesial-temporal region (Manns & Squire, 2002).

Long-term verbal memory is measured with the delayed free recall, cued recall, and recognition trials of tests. In the current study, Logical Memory-2 was compared with the Delayed Recall scores of the HVLT-R, and the percentage retention rates on the Logical Memory subtest were compared with the percentage retention rates on the HVLT-R.

As Table 10 shows, the delayed recall trial of the Logical Memory subtest (LM-2) correlated higher with the Delayed Recall trial of the HVLT-R (H-DR) (r = .658) than any other HVLT-R task This should be expected due to the two measures (LM-2 and H-DR) being comparable tasks. Specifically, these are both tasks that tap the mesial-temporal region. Likewise, the highest correlation between the percentage retention component of the Logical Memory subtest (LM-PR) and any of the HVLT-R scores was with the percentage retention category of the HVLT-R (H-PR) (r = .519).

As with the comparison between the immediate memory measures, exploratory analysis was conducted in order to determine the frequency of each delayed verbal memory task's categorization of subjects from the acute mild head-injured sample as either experiencing memory difficulty to the point of placement in the impaired range or not. The raw scores for LM-2 were compared with norms from the test manual (Wechsler, 1987), and percentiles for the subjects were obtained. Likewise, the raw scores from H-DR, and H-PR were compared with normative data found in the test's manual (Brandt & Benedict, 2001), and T-scores for the subjects were obtained. As with the immediate memory comparisons, a standard criterion of 2 standard deviations below the mean to qualified as an impaired score.

Of the 698 subjects with complete HVLT-R scores, 56 were not administered Logical Memory-2 subtest. Similar to the comparison between H-1, H-TR, and LM-1, a comparison was also made between the delayed recall components of the HVLT-R (H-DR) and Logical Memory (LM-2). The results can be found in Table 21. As with the comparison on the initial trials, these results indicate that subjects in the current study with a diagnosis of status post mild closed head injury were more likely to be assessed as having memory impairment when administered the HVLT-R than the Logical Memory -2 subtest of the WMS-R.

Table 21

Frequency of Impairment Between H-DR and LM-2 (N = 642)

Age Group	H-DR	LM-2
18-19 years old (N = 49)	27	9
20-29 years old (N = 152)	92	24
30-39 years old (N = 112)	54	13
40-49 years old (N = 127)	63	10
50-59 years old (N = 97)	69	13
60-69 years old (N = 37)	21	3
70-79 years old (N = 46)	25	2
80+ years old (N = 22)	11	7
TOTAL	362	81

Overall, H-DR identified 362 subjects (56.4% of the total subjects compared) as having some level of impairment in memory. In contrast, the LM-2 subtest identified 81 subjects (12.6%) as experiencing memory impairment. Table 22 provides a breakdown of the percentage of subjects identified as impaired in each age group. As Table 22 shows, the Delayed Recall score of the HVLT-R identified a higher percentage of subjects in each age group as experiencing memory impairment when compared with the performance of the LM-2.

Table 22

Percentage of Subjects Per Age Group Identified as Impaired Between H-DR and LM-2 (N = 642)

Age Group	H-DR	LM-2
18-19 years old (N = 49)	55.1	18.4
20-29 years old (N = 152)	60.5	15.8
30-39 years old (N = 112)	48.2	11.6
40-49 years old (N = 127)	49.6	07.9
50-59 years old (N = 97)	71.1	13.4
60-69 years old (N = 37)	56.8	08.1
70-79 years old (N = 46)	54.4	04.4
80+ years old (N = 22)	50.0	31.8
TOTAL	56.4	12.6

The HVLT-R professional manual (Brandt & Benedict, 2001) provides normative data for interpretation of the percentage that subjects retain information from the highest raw score of the initial trials (trial 2 or 3) to the delayed recall trial. The manual for the WMS-R (Wechsler, 1987), however, does not provide normative data for interpreting percentage retention on the Logical Memory subtest. In order to compare the percentage retention rates between the HVLT-R and the Logical Memory subtests, the cutoff percentages on the HVLT-R that differentiate between intact and impaired performance were used with the current sample for the percentage retention on both the HVLT-R and the Logical Memory. The cutoff percentages were stratified by age group. Results can be found in Table 23. As Table 23 indicates, overall H-PR rates were higher than LM-PR rates, and this pattern held for each age group. However, the differences between the percentage retention on the HVLT-R scores (overall 47.9%) and the LM scores (41.3%) appeared to be rather small. Several of the age groups also showed very little difference between the two measures. This consistency in performance across measures is likely due to both tapping mesial-temporal functioning.

Table 23

Frequency and Percentage of Impairment on HVLT-R Percentage Retention and Logical Memory Percentage Retention

Age Group	H-PR	H-PR Percentage	LM-PR	LM-PR Percentage
18-19 years old	20 (N = 54)	37.0	18 (N = 49)	36.7
20-29 years old	84 (N = 163)	51.5	75 (N = 152)	49.3
30-39 years old	48 (N = 121)	39.7	35 (N = 112)	31.3
40-49 years old	59 (N = 140)	42.1	43 (N = 128)	33.6
50-59 years old	60 (N = 99)	60.6	53 (N = 97)	54.6
60-69 years old	19 (N = 40)	47.5	15 (N = 37)	40.5
70-79 years old	29 (N = 51)	56.9	14 (N = 46)	30.4
80-89 years old	15 (N = 30)	50.5	10 (N = 22)	45.5
TOTAL	334 (N = 698)	47.9	263 (N = 636)	41.3

Relationship between the LM Scores, HVLT-R Scores and COWAT

The COWAT is included in the neuropsychological battery because of its link to the frontal cortex and, more specifically the left frontal cortex. This executive functioning measure of word fluency and speech has been found to be sensitive to left frontal and bilateral frontal lesions. Damage to the frontal lobes is a common outcome in acceleration-deceleration injuries and diffuse axonal injuries (Sheid, Preul, Gruber, Wiggins, & Cramon, 2003).

Significant correlation coefficients ($p < .01$) between the COWAT and the LM-1, LM-2, H-1, H-2, H-3, H-TR, H-DR, and H-PR were found and are provided in Table 24. As these results indicate, the correlation between each of the HVLT-R tasks is higher than each of the Logical Memory subtests, with the exception of Percentage Retention on the HVLT-R.

Table 24

Correlation Coefficients between COWAT and Scores on the HVLT-R, LM-1 and LM-2

Measures versus COWAT	r
HVLT-R Total Recall	.450
HVLT-R Trial 1	.393
HVLT-R Delayed Recall	.379
Logical Memory Trial 2	.328
Logical Memory Trial 1	.325
HVLT-R Percentage Retention	.278
Logical Memory Percentage Retention	.223

The Total Recall score of the HVLT-R ($r = .450$) and the first trial of the HVLT-R ($r = .393$) both were higher than Logical Memory 1 ($r = .325$). Fisher's Z transformation was calculated for these correlation coefficients to determine if they significantly differed. The correlation between the COWAT and the Total Recall of the HVLT-R was significantly higher than the correlation between the COWAT and Logical Memory-1 ($z = 2.608$; $p < .01$). The correlation between the COWAT and the first trial of the HVLT-R did not significantly differ from the correlation between the COWAT and Logical Memory-1 ($z = 1.268$; $p > .05$).

On the delayed recall tasks, the correlation between the COWAT and the Delayed Recall score of the HVLT-R ($r = .379$) was higher than the correlation between the COWAT and Logical Memory 2 ($r = .328$). However, Fisher's z transformations found that the difference in these correlations was not significant ($z = 1.053$; $p > .05$). The correlation between the COWAT and HVLT-R percentage retention ($r = .278$) was not significantly higher than the correlation between the COWAT and Logical Memory percentage retention ($r = .223$) ($z = 1.77$; $p > .05$). Although these differences appear to be marginal, a pattern emerges in which correlations with the HVLT-R compared to Logical Memory was higher each time, even if not significantly higher. This consistent pattern may be due to the two verbal memory tests measuring different aspects of verbal memory or perhaps tapping different brain regions.

With the general standard of two standard deviations below the mean (Spreen & Strauss, 1998) as the cutoff, exploratory data analysis was conducted to determine the number of subjects who scored in the impaired range on the COWAT and each of the memory scores from the HVLT-R and Logical Memory subtests. The results, broken down by age group can be found in Table 25. To be included, subjects had to score in the impaired range on the COWAT and at least one of the memory measures. For example, the 5 subjects under the LM-1 column and 18-19-year row had to score in the impaired range on the COWAT and LM-1 to be counted.

As Table 25 indicates, of the 678 subjects administered the COWAT, scores on the measure identified 140 subjects (20.65%) as being in the impaired range. Of the memory tasks, the Delayed Recall of the HVLT-R identified 116 of those subjects (82.9%) as scoring in the impaired range. The Total Recall trial of the HVLT-R identified 107 of the subjects (76.4%) as impaired. In contrast, the percentage of subjects scoring in the impaired range on the COWAT and either Logical Memory 1 (20.0%) or Logical Memory 2 (23.8%) was considerably lower than the percentages of individuals scoring in the impaired range on the COWAT and the HVLT-R scores. As with the data in Table 24, these results may be due to the two memory tests measuring different aspects of verbal memory or even different brain regions.

Table 25

Frequency and Percentages of Subjects per Age Group on Memory Task in Agreement with Impaired COWAT

Age Group	LM-1 (n = 626)	LM-2 (n = 625)	H-TR (n = 678)	H-DR (n = 678)	H-PR (n = 678)	COWAT (n = 678)
18-19 years old	5 (55.6%)	4 (44.5%)	8 (88.9%)	9 (100.0%)	4 (44.5%)	9 (100%)
20-29 years old	8 (26.7%)	11 (36.7)	18 (60%)	26 (86.7%)	22 (73.3%)	30 (100%)
30-39 years old	3 (15.0%)	3 (15.0%)	17 (85.0%)	18 (90.0%)	12 (60.0%)	20 (100%)
40-49 years old	1 (4.5%)	4 (18.2%)	20 (90.9%)	19 (86.4%)	16 (72.7%)	22 (100%)
50-59 years old	3 (13.6%)	5 (22.7%)	17 (77.3%)	18 (81.8%)	15 (68.2%)	22 (100%)
60-69 years old	2 (25.0%)	0 (0.0%)	5 (62.5%)	7 (87.5%)	7 (87.5)	8 (100%)
70-79 years old	3 (18.8%)	2 (12.5%)	12 (75.0%)	11 (68.8%)	11 (68.8%)	16 (100%)
80+ years old	3 (23.1%)	4 (30.8%)	10 (76.9%)	8 (61.5%)	8 (61.5%)	13 (100%)
TOTAL	28 (20.0%)	33 (23.8%)	107 (76.4%)	116 (82.9%)	95 (67.9%)	140 (100%)

Relationship between the LM Scores, HVLT-R Scores and Trails B

Part B of the Trail Making Test has also been found to be sensitive to brain damage and, in particular, frontal lobe dysfunction. Of the 698 subjects in the current study, only 340 were administered the Trails B task. The majority of the subjects excluded from this analysis did not have scores for Trails B due to skeletal injuries which precluded them from manipulating a pencil. Each of the 340 subjects were administered the HVLT-R and Trails B, and 320 of the subjects were also administered the Logical Memory subtests.

Significant correlation coefficients ($p < .01$) between Trails B and the LM-1, LM-2, LM-PR, H-1, H-2, H-3, H-TR, H-DR, and H-PR were found and are provided in Table 26. As these results indicate, the correlations with each of the HVLT-R tasks is higher than each of the Logical Memory subtests, with the exception of Percentage Retention on the HVLT-R. Nevertheless, H-PR is still more strongly correlated with Trails B than LM-PR.

Fisher's z transformations found that the correlation between Trails B and the Total Recall of the HVLT-R ($r = .549$) was significantly higher than the correlation between Trails B and Logical Memory-1 ($r = .401$) ($z = 2.429$; $p < .01$). The first trial of the HVLT-R did not correlate significantly higher with Trails B than did Logical Memory-1 ($z = 1.079$; $p > .05$).

On the delayed recall and percentage retention measures, no significant differences were found between the correlations. The correlation between Trails B and Delayed Recall of the HVLT-R did not significantly differ from Trails B and Logical Memory-2 ($z = 0.150$; $p > .05$). Likewise, the correlation between HVLT-

R Percentage Retention and Trails B did not significantly differ from the correlation between Logical Memory Percentage Retention and Trails B (z =-0.405; p = ns).

Table 26

Correlation Coefficients between Trails-B and Scores on the HVLT-R, LM-1 and LM-2

Measures Versus Trails B	r
HVLT-R Total Recall	-.549
HVLT-R Trial 1	-.472
HVLT-R Delayed Recall	-.412
Logical Memory Trial 2	-.408
Logical Memory Trial 1	-.401
HVLT-R Percentage Retention	-.293
Logical Memory Percentage Retention	-.255

Summary of Findings

The current study found that, on measures designed to tap immediate verbal memory, the HVLT-R and Logical Memory subtest differed substantially in the number of subjects each placed in the impaired range. Similarly, on the delayed recall measures of the HVLT-R and Logical Memory, a substantial difference in subjects placed in the impaired range was also found. A comparison of the percentage retention for the two tasks, however, did not result in a significantly noticeable difference.

The sizable differences between the HVLT-R and Logical Memory on immediate and delayed recall are in terms of the number of people that the HVLT-R identifies as being impaired. It is possible that the HVLT-R is simply an overall more difficult task than the Logical Memory subtests. However, it is also possible that mild-TBI involving frontal lobe dysfunction makes certain tasks, such as word lists, more difficult for subjects to master than other measures. In order to address this issue groups were created that reflect presumptive frontal lobe dysfunction and those without dysfunction.

Exploratory Analysis of Frontal Lobe Superior versus Impaired

Two groups, stratified by age group, were constructed from the original 698 subjects. Inclusion in this analysis required subjects to have complete scores on the HVLT-R, Logical Memory subtests, COWAT and Trails B. Three hundred fourteen subjects were retained.

Of the 314 subjects, a “frontal lobe superior” subsample was created based on subjects who scored in the top quartile of both the COWAT and Trails B, two neuropsychological tasks that purport to tap frontal lobe functioning. Likewise, a “fontal lobe impaired” subsample was created that included subjects who had scores in the bottom quartile on both the COWAT and Trails B. This resulted in 42 subjects in the frontal lobe superior range and 43 subjects in the frontal lobe impaired range. Table 27 shows the raw score means and standard deviations for the two groups.

Independent-sample t-tests were conducted for each memory task. Table 27 provides the results of the t-tests as well as the effect sizes. As Table 27 indicates, the superior group scored significantly higher than the impaired group on each verbal memory task. Effect sizes for all the measures were high, except for percentage retention on both the HVLT-R and Logical Memory. The results of this analysis indicates that the HVLT-R and Logical Memory subtests both distinguish between subjects with intact frontal lobe functioning and subjects with impaired frontal lobes. The comparison of means (t-tests) between the groups revealed that the frontal lobe superior group performed significantly better than the frontal lobe impaired group on each verbal memory task. Therefore, the results of the independent t-tests do not support the argument that the HVLT-R is more affected than Logical Memory by frontal lobe dysfunction.

Table 27

Means, SD, T-Tests and Effect Size for Frontal Lobe Superior (FLS) and Impaired (FLI) Groups (COWAT and Trails B)

	Frontal Lobe Superior		Frontal Lobe Inferior			
Memory Test	M	SD	M	SD	t	Effect Size (r)
LM1	22.6	6.8	15.1	6.8	5.05**	48
LM2	17.7	8.0	9.1	5.7	5.70**	.53
LM-PR	76.7	20.2	58.4	29.2	3.37**	.34
H1	6.2	1.1	4.1	1.6	6.88**	.61
H-TR	23.9	4.3	15.4	4.9	8.49**	.68
H-DR	7.9	3.1	3.7	2.7	6.55**	.59
H-PR	81.1	24.1	55.5	39.6	3.60**	.37

** Equals p value less than or equal to .001

In order to examine a larger contingent of the overall subjects in this study, the COWAT alone was utilized to create two groups. Of the 626 subjects who completed the COWAT, the top and bottom quartile was extracted resulting in 157 frontal lobe superior subjects and 153 frontal lobe impaired subjects. Table 28 shows the raw score means and standard deviations for the two groups.

Independent-sample t-tests were conducted for each memory task. Table 28 provides the results of the t-tests as well as the effect size. As Table 28 indicates, the frontal lobe superior group significantly outperformed the frontal lobe impaired group on each verbal memory task.

The analyses of the frontal lobe superior and frontal lobe impaired groups were similar when the COWAT was used to establish the groups and when the COWAT was used in conjunction with Trails B to establish the groups. Both analyses revealed that, when the frontal lobes are impaired, the Logical Memory subtests are just as successful as the HVLT-R at identifying the subjects as brain injured. Therefore, in the current study the HVLT-R did not prove to be a superior measure of frontal lobe functioning.

Table 28

Means, SD, T-Tests and Effect Size for Frontal Lobe Superior (FLS) and Impaired (FLI) Groups (COWAT only)

	Frontal Lobe Superior		Frontal Lobe Impaired			
Memory Test	M	SD	M	SD	t	Effect Size (r)
LM1	20.9	7.0	13.9	6.3	9.25**	.47
LM2	15.9	7.8	8.7	5.9	9.18**	.48
LM-PR	73.7	22.5	58.1	32.4	4.90**	.29
H1	5.8	1.4	4.1	1.4	10.75**	.52
H-TR	22.4	4.5	15.6	4.9	12.70**	.59
H-DR	6.8	3.1	3.3	3.0	10.06**	.50
H-PR	72.8	27.9	45.2	37.6	7.34**	.40

** Equals p value less than or equal to .001

CHAPTER V: DISCUSSION

Mild-TBI is the most common form of brain trauma. Kraus and Chu (2005) estimated that 50% of all hospital admissions in the United States due to brain injury are mild traumatic brain injuries. Many mild-TBI cases may go unreported because the individual suffering the TBI refuses or otherwise fails to seek medical assessment and treatment. Of the mild-TBI cases that do seek medical assistance, Kraus and Chu (2005) estimated that 100% are eventually discharged from the hospital. However, the use of the word “mild” to describe these injuries may be misleading, because there certainly are individuals who suffer a mild-TBI that subsequently results in significant impairment in cognitive functioning. Because of the vital need to be able to distinguish between patients who are having difficulty in cognitive functioning and those who are not, neuropsychologists need to have access to reliable and valid tests that will assist them in making the distinction between impaired and intact. The current study examined the use of the HVLT-R as a neuropsychological screening tool of verbal memory with a large sample of mild-TBI subjects.

Interpretation of Findings

This study looked specifically at whether differences in age and education affected performance on the Total Recall, Delayed Recall, and Percentage Retention measures of the HVLT-R with mild TBI subjects. A second intent of the study was to consider whether the HVLT-R tasks perhaps tap different aspects of the brain’s memory system compared to other verbal memory tasks, specifically the Logical Memory subtests of the WMS-R. A final intention of the current study was to examine whether the first trial of the HVLT-R compared favorably to another verbal memory task often utilized as a measure of attentional ability, namely the Digit Span subtest of the WMS-R.

Although the HVLT has been in existence since 1991 and its revised version, the HVLT-R, has been utilized since 1998, very little research has been conducted with this test as a useful screening measure for verbal memory in a head injured population. The bulk of the HVLT-R studies have targeted dementia and stroke, focusing particularly on the elderly. The importance of the current study is that it focuses on a very large sample of mild head injured subjects and includes subjects from all adult ages, rather than targeting primarily the elderly.

Hypothesis 1: Effects of Age on HVLT-R Performance with Mild-TBI

Regarding the individual hypotheses, it was hypothesized that a difference in performance on the HVLT-R measures would be found based on age. This hypothesis was supported. Age differences were found on each of the HVLT-R measures investigated. Groups consisting of individuals with mild-TBI were found to have less restrictive age differences than the HVLT-R professional manual (Brandt & Benedict, 2001) suggests for the normative group. While the professional manual groups age in 10-year increments, the current study found that many of these 10-year age groups did not significantly differ from each other, and therefore evidence suggests that the 8 age groups could be reduced to 3 age groups: 18-29 years, 30-69 years, and 70+ years. This finding is consistent with normative data introduced by Benedict and his colleagues (1998) whose data grouped HVLT-R subjects into 4 age groups. However, normative data for similar word-list verbal memory tasks, such as the Rey Auditory Verbal Learning Test (RAVLT; norms in Spreen & Strauss, 1998) are often grouped by age in 10 year increments very similar to the HVLT-R Professional Manual.

Although normative data for the HVLT-R has been published for special groups (Friedman et al. 2002; Vanderploeg, et al. 2000), only two sets of HVLT-R norms have set standardization data from late teen years through late adulthood. These two sets of norms, Benedict et al. (1998) and Brandt et al. (2001) are essentially in disagreement as to how narrow the age ranges need to be. The current findings have shown that, with a large sample of individuals with mild traumatic brain injury, restrictive age group ranges are not necessary and supports the findings of Benedict et al. (1998).

Hypothesis 2: Effects of Education on HVLT-R Performance with Mild-TBI

It was predicted that significant differences would be found on HVLT-R performance based on level of education. This hypothesis was confirmed. Significant differences in performance on HVLT-R measures due to level of education were found. The current study found that, although mild-TBI patients who have more than a high school education differed significantly from those with a high school education or less, there was no significant difference between individuals whose academic experience ended after graduating from high school and those who did not complete high school. Therefore, only two groups appear necessary to differentiate the impact that education level has on HVLT-R performance: those with a high school education or less, and those with more than a high school education. The effect that education has on performance varies with each neuropsychological test. For example, Crum, Anthony, Bassett, & Folstein (1993) established normative data for the Mini-Mental State Examination (Folstein et al., 1975) that closely reflected the education groups found in the current study. Crum and colleagues established education categories of 0-4 years of education, 5-8 years, 9-12 years (including high school diploma) and "college experience or higher degree." Norms for the COWAT and Animal Naming (see Ruff et al., 1996; Tombaugh, Kozak, & Rees, 1996; each cited in Spreen & Strauss, 1998) tend to be stratified similarly.

Of the five sources of normative data for the HVLT or HVLT-R, there has been disagreement concerning whether or not level of education has an effect on performance. Although Benedict et al. (1998) reported the average highest level of education from their sample, neither they nor Brandt (1991) considered level of education in the actual standardization data. Vanderploeg, et al. (2000) investigated the effect of education with their standardization sample, but concluded that level of education did not have an effect on HVLT-R performance. Friedman, et al. (2002) confirmed education effects and incorporated an education-corrective component to their norms. However, the normative data provided by Friedman and colleagues was exclusively designed for African Americans and, therefore, use of their norms would be inappropriate with a majority of patients. Brandt and Benedict (2002) recognized that education level had an effect on HVLT-R performance, but nevertheless failed to account for level of education in the stratification of their normative data. The current findings, however, support the need for development of HVLT-R norms that are stratified, not only by age, but also by level of education.

Hypothesis 3: HVLT-R and Logical Memory

It was predicted that the HVLT-R measures would not strongly correlate with the measures of another verbal memory test, the Logical Memory subtests of the WMS-R. This hypothesis was confirmed. Modest correlations were found between equivalent measures on the HVLT-R and Logical Memory. However, the impetus for this research question was that, if confirmed, the results may provide evidence for the belief that verbal word list tests tap a different aspect of verbal memory functioning than prose recall. In particular, Tremont and colleagues (2000) stated that clinical experience pointed to patients with executive dysfunction performing considerably worse on list-learning tasks than on prose-recall tasks. They concluded that the results of their study confirmed their clinical observations. However, based on exploratory analysis, the current study did not support the conclusion that subjects with impairment in executive functions perform better on prose-recall than list-learning tasks.

Although each of the HVLT-R tasks placed substantially more subjects in the impaired range when compared to the corresponding Logical Memory task, there was no significant difference in the ability of the HVLT-R tasks compared to the Logical Memory tasks to distinguish between frontal lobe intact groups and frontal lobe impaired groups. Independent t-tests found that the frontal lobe inferior group performed significantly worse than the frontal lobe superior group on the first trial, Total Recall, Delayed Recall, and Percentage Retention of the HVLT-R. Likewise, independent t-tests showed that the frontal lobe inferior group scored significantly worse than the frontal lobe superior group on Logical Memory-I, Logical Memory-II, and Logical Memory Percentage Retention. In other words, each HVLT-R task and each Logical Memory task were able to differentiate between the two frontal lobe groups because of the significantly poorer performance of the frontal lobe inferior group on each of the verbal memory tasks.

Of the 698 subjects in the current study, the Logical Memory-1 subtest only identified 8.9% of the subjects as experiencing impairment in immediate verbal memory. This percentage is likely quite low considering that memory difficulty is one of the most frequent results of a closed head injury (Levin, Lilly, Papanicolaou, & Eisenberg, 1992). In contrast to the performance of the Logical Memory-1 subtest, the first trial of the HVLT-R and the Total Recall score both identified 33.1% and 46.7% of the subjects respectively as experiencing impairment in their immediate auditory memory functioning.

Findings on the delayed recall tasks were similar to those found on the immediate recall tasks. Of the 698 subjects, Logical Memory-2 only identified 12.6% of the subjects as experiencing impairment in delayed verbal memory. In contrast, the H-DR identified 56.4% of the subjects. In contrast, the percentage retention rates for both Logical Memory and HVLT-R were relatively similar.

When comparing the HVLT-R measures and the Logical Memory subtests with the frontal lobe superior and frontal lobe impaired groups created, the effect sizes for each HVLT-R measure and the corresponding Logical Memory measure were relatively similar. Although there are several potential explanations, one possible explanation for this finding is that the HVLT-R is a more difficult task than the Logical Memory subtests and is, therefore, more likely to assess a mild-TBI patient as impaired in verbal memory as a result of test difficulty, rather than the tapping of different brain regions.

Although the HVLT-R was originally intended to be used as a verbal memory screening device when more complex word lists were inappropriate due to the severity of the patient’s impairment, the test may still make certain demands on mildly head injured patients that they find difficult. For example, the majority of the verbal word list tests currently in use by neuropsychologists only contain words that are unrelated to other words on the list. In contrast, the HVLT-R groups words into three semantic categories. This may make it easier for an intact individual to master the word list because of the use of organizational skills and the ability to chunk the words into semantic categories (Lezak 1983). However, for an individual with mild-TBI, the ability to use chunking and other organizational skills may be impaired. Therefore, the advantage of the HVLT-R for intact individuals becomes a disadvantage for the impaired individual. This may explain the very large number of subjects in the current study who performed in the impaired range on at least one of the HVLT-R measures.

The Logical Memory subtests are likely easier to master because points are awarded for close approximations to certain aspects of the story. For example, Story A of Logical Memory-1 has the phrase “of fifty-six dollars.” However, subjects are awarded full credit for indicating any dollar amount between 49 and 60 dollars. In contrast, subjects are not awarded points on word list tests such as the HVLT-R unless the words they remember are the exact ones on the list. Differences, therefore, between the HVLT-R and Logical Memory may be due to the differences in expectations of precise recall of the information. Not only do the Logical Memory subtests award points for getting the gist of the story, individuals may also find the contextual presentation of prose memory tasks makes it easier to encode and store the verbal information for later retrieval. The HVLT-R does not offer these advantages.

Ecological Validity of the Current Findings

Sbordone (1997) defined ecological validity as “the functional and predictive relationship between the patient’s performance on a set of neuropsychological tests and behavior in a variety of real-world settings.” (p. 368). The ecological validity concern is the question of transferring the knowledge of the patient’s test scores in a sterile, quiet, distraction-free environment to a prediction of how this specific patient will adapt and function in his or her daily life outside of the laboratory, hospital or neuropsychologist’s office (in the real-world setting).

Long (1998) noted that referrals for neuropsychological assessment frequently request that neuropsychologists make predictions concerning the patient's functional consequences subsequent to brain damage. This involves assessing the patient's strengths and weaknesses in several cognitive domains. Sbordone and Long (1998) stated that neuropsychologists working in rehabilitation settings are also asked to give their opinions not only concerning the patient's functional ability but also to estimate rehabilitation potential, recommend treatment options, and recommend optimal living arrangements. The opinions of the neuropsychologist are important because they may impact the patient's potential for certain educational and occupational opportunities as well as the patient's ability to manage their own personal affairs (Long, 1998).

Rarely is a neuropsychological test developed with the intent of tapping the functioning of a specific cognitive domain in the patient's real-world environment. The Rivermead Behavioural Memory Test (RBMT), developed by Wilson, Cockburn and Baddeley (1985), is one example of a memory test designed to predict behavior in a real-world environment and provide information concerning a patient's everyday problems with memory. However, the majority of neuropsychological measures were designed to assess organic impairment and the use of many of these tests in predicting behavioral outcome is still in need of further investigation.

Franzen and Wilhelm (1998) asserted that there are two aspects of ecological validity. They termed the ability of a neuropsychological test's results to predict real-world functioning as "veridicality." The second aspect is “verisimilitude”, which they define as “the similarity of the data collection method to tasks and skills required in the free and open environment.” (p. 93). According to Franzen and Wilhelm verisimilitude may be the more important of the two aspects in terms of neuropsychological test design. Unfortunately, according to Goldstein (1998), the attainment of verisimilitude is virtually impossible because it would be impossible to design neuropsychological tests intended to be administered in the natural setting that the behavior in need of assessment occurs. Even if neuropsychological tests were designed to be administered in the real-world settings (for example, in the patient’s home), the patient may behave differently than they normally would in that environment because they know that they are being observed by the evaluator.

With the above caveats in mind, it is important to consider the implications of the current study’s findings on the association between brain functioning and behavioral outcome. A high percentage of subjects in the current study performed in the impaired range on the various HVLT-R tasks, including the first trial (33%), Total Recall score (47%), Delayed Recall score (56%) and percentage retention (48%). In contrast, a much lower percentage of subjects scored in the impaired range on Logical Memory I (9%), Logical Memory II (13%), and the percentage retained from Logical Memory I to Logical Memory II (41%). Discussion of the verisimilitude as it pertains to these measures is speculative because traditional verbal memory tasks such as these were not designed with application to real-world settings in mind. Nevertheless, certain distinctions can be made.

As already mentioned above, the Logical Memory tasks require patients to focus on retaining and recalling the gist of two paragraph-length stories. Certain words contained in the story must be repeated verbatim in order for points to be awarded. However, for other words or phrases found in the story, points are awarded for close approximations to the words that were actually read by the evaluator. In real world settings, there are certain instances in which precise memorization of verbal conversations, stories, requests and commands are not necessary and, even for most non-injured individuals, a word-for-word recitation would not be possible. What is necessary in those instances is that the individual is able to comprehend the general content of the verbal information being spoken.

There are numerous other situations, however, in which the precise capability to store and retrieve verbal information is vital. One very important example of this type of environment is the academic setting. Many academic exercises require rote memorization of facts and information. The employment setting is another example. Most vocations require the ability to precisely memorize information to some degree. When an individual returns to work post-injury, he or she may exhibit limited cognitive abilities that can impede performance of one's job responsibilities.

Memory Impairment Versus Impaired Memory Score

It is necessary to first distinguish between impaired behavior and impaired performance on a neuropsychological measure. For the individual subject who scores in the impaired range, this simply means that, on a given verbal memory task, his or her performance and associated score was, when compared to normative data of healthy peers, so low as to be considered abnormally low. It does not, however, necessarily imply that the individual actually has impaired verbal memory (immediate or delayed). It is worth considering Sbordone's (1997, 1998) warnings of false assumptions that handicap the neuropsychologist, and in particular the assumption that defective performance on neuropsychological tests indicates cognitive impairment.

When performance on an individual memory test score is far below the norm, it can act as a warning to the neuropsychologist that further evaluation, follow-up, or specific recommendations may be warranted. However, the manifestation of impaired memory functioning is more appropriately found in the patient's behavior in his or her day-to-day real-world environments including home, school, and work rather than a score in the impaired range on a memory test administered in the sterile confines of an inpatient or outpatient neurological treatment setting.

Interpretations of Discrepancy Between Logical Memory and the HVLT-R

The results of the current study did not support the conclusions by Tremont and colleagues (2000) concerning the connection between frontal lobe dysfunction and poorer performance on list-learning tasks than on story-recall tasks. Independent t-tests confirmed that a frontal lobe inferior group performed significantly worse than a frontal lobe superior group on each HVLT-R and

Logical Memory task. Therefore, other potential reasons must be explored to account for the substantial difference in the percentage of subjects scoring in the impaired range on HVLT-R tasks versus Logical Memory tasks.

One possibility for the discrepancy is that the HVLT-R, instead of being a more sensitive measure to frontal lobe functioning, may instead be more sensitive to global neurological dysfunction. Indeed, this is the conclusion that Busch et al. (2005) derived in a study also designed to compare the performance of another verbal word list, the CVLT, with the Logical Memory subtest in correlating with frontal lobe tasks. In fact, Busch and colleagues (2005) asserted that the tapping of global dysfunction may be a more accurate interpretation of the Tremont and colleagues (2000) study.

Another possibility for the high number of impaired scores on the HVLT-R measures versus the Logical Memory measures is the possibility of a difference in test difficulty between the two measures. Tremont and associates (2000) offered this as one alternate conclusion to their findings as well. Guilmette and Rasile's (1995) findings were similar to the current study. They compared the Logical Memory subtests with the Rey Auditory Verbal Learning Test (RAVLT), a word-list verbal memory task similar to the HVLT-R and the Expanded Paired Associates Test (EPAT). The study included 8 men and 8 women who had suffered a mild brain injury and were subsequently referred for an outpatient neuropsychological evaluation, and a control group made up of 8 men and women volunteers from the community. Subjects were administered three verbal learning and memory measures as a part of a neuropsychological battery. The learning and memory tasks consisted of the Logical Memory subtests of the WMS-R, the RAVLT, EPAT. Paired t-test were analyzed and the results demonstrated that the mild brain injured group performed significantly worse than the control group on the majority of the memory measures. A hierarchy of difficulty was also found with the EPAT being the most difficult verbal memory task, followed by the RAVLT, and then the Logical Memory subtests. This pattern was consistent for the scores of both the mild brain injured group and the control group and the pattern held throughout all tasks. In other words, with a comparison of the converted standard scores for each measure, the both of the EPAT scores for both groups were lower than all of the scores for the RAVLT and Logical Memory. Likewise, all of the RAVLT standard scores for both groups were lower than the scores on Logical Memory I and II. This finding indicates that a word list memory task may be a more difficult verbal memory task than a prose memory task, not only for brain injured individuals but also for healthy individuals.

The Guilmette and Rasile (1995) findings are significant in terms of the clinical utility of these memory measures. For the mild brain injured group, the average scores for Logical Memory I and II were in the average range (37^{th} percentile). This meant that many of the mild brain injured subjects were not scoring in what is the standard impaired range of below the 3^{rd} percentile. The authors suggested that, if sensitivity (i.e., true positives) is a concern for the neuropsychologist using the Logical Memory subtests to assess a mild brain injured patient, the cutoff scores for impairment may need to be adjusted (increased) considerably. They found that, when assessing for mild brain injury, the highest level of diagnostic accuracy for both the Logical Memory I and II subtests was found when the impairment cutoff score was at or below a standard score of 95 (37th percentile). In contrast, the optimal cutoff score for detecting mild brain injury with the RAVLT was at or below a standard score of 85 (16th percentile) with the lone exception being the Long delay score (standard score at or below 78). The authors concluded that, when using the verbal memory tests to assess mild brain injury, the neuropsychologist must decide whether sensitivity (i.e., true positives) or specificity (i.e., true negatives) is more important and this decision may be based on the purpose of the evaluation. If sensitivity is of greatest concern, impairment cutoff scores may need to be increased especially for the Logical Memory subtests.

The HVLT-R contains words that are organized into three semantic categories; however, on each of the alternate forms of the HVLT-R, it is a rare occurrence for a word from one semantic category to be presented immediately following a word from the same category. Therefore, if the subject is going to benefit from the clustering of semantically-related words, he or she will be required to exhibit the organizational skills to do so. Many post-trauma patients are unable to utilize such organizational skills. The Logical Memory subtest, on the other hand, provides a contextually-rich story which may make memorization easier. The story is already organized, so rearranging the content in order for it to make sense is not necessary. Finally, as has been previously stated, points are awarded on the Logical Memory subtests for close approximations to the content of the story. The HVLT-R does not allow points to be awarded for close approximations.

Still other possibilities include the effect that an alteration in personality due to a brain injury can have that would impact performance or effort. For example, Marin and Chakravorty (2005) noted that loss of motivational factors is often affected by traumatic brain injury because motivation is impacted by decreased dopaminergic activity due to damage to key brain regions including the hippocampus and amygdala in the mesial-temporal region as well as the prefrontal cortex. The prefrontal cortex, hippocampus, and amygdala modulate the drive state of the individual. Significant damage to these brain regions may result in akinetic mutism, abulia, or apathy. Marin and Chakravorty asserted that lack of motivation may be due to the damage to mesial-temporal and dorsolateral structures, and also may be due to a psychological response to the individual's perception that their efforts at a task will not be successful due to their lack of ability to organize behavior. In the current study and similar studies that compare prose recall to word list recall (for example: Tremont et al, 2000; Busch et al., 2005), subjects may find the story recall task to be less intimidating because the story is set in a rich context and is already organized for the subject. In contrast, a verbal word list does not provide contextual cues, and as previously stated, requires the subject to possess and exhibit organizational skills if he or she is to benefit from the semantic clustering that could aid their recall. With lesions to the frontal and mesial-temporal regions during a mild-TBI, the subject may need less motivation to complete the contextual, organized task (Logical Memory), perhaps because recalling a short story may seem less threatening than a word list or perhaps because the continuity of a story may spark some interest in the subject, increasing motivation to listen and recall. A verbal word list may seem more intimidating and lacks the interest component of a story. Therefore, the subject may lack the ability to generate enough motivation to complete the more complex verbal memory tasks found in word lists. Lezak (1995) noted that subjects who are incapable of learning 10 to 15 words view word list tests with embarrassment and/or drudgery and may not see any practical purpose in memorizing a list of words. Therefore, low face validity in word list tasks may also affect motivation.

Another emotional characteristic that could potentially influence the results of the current study is alteration in emotional capacity. Although not every depressed individual experiences difficulty with memory, depression is one of several residuals secondary to traumatic brain injury (Culum, Kuck, & Ruff, 1990). In a review of several studies published between 1981 and 1991 of the depressive symptom prevalence in traumatic brain injury, Robinson and Jorge (2005) found that estimates of depression associated with traumatic brain injury to be between 25% and 50% including an estimate for depressive symptoms in patients with mild brain injury at 39%. In a study investigating the relationship between verbal memory and depression following traumatic brain injury, Keiska, Shore and Hamilton (2007) found an association between depression and diminished performance on delayed recall and recognition of the California Verbal Learning Test-II (CAVLT-II), but did not find a significant relationship between depression and recall on either Logical Memory or Verbal Pairs Associates subtests of the WMS-III. Lezak (1995) discussed the effects that damage to the frontal lobes and the neural pathways connecting the cerebral cortex with the diencephalon which contains the centers for integrating affect and drive. She noted that there is a connection between loss of affective capacity and low motivation. Although limitations in the current study preclude the exploration of the extent alteration in emotional capacity may have had on the results, the results may have been impacted by a loss of affective capacity and subsequent lower drive states in the frontal lobe inferior group.

Hypothesis 4: HVLT-R Trial 1 and Digit Span

It was hypothesized that the first trial of the HVLT-R would adequately correlate with the Digit Span subtest of the WMS-R, suggesting that the Trial 1 of the HVLT-R could be utilized by clinicians as a tool for the assessment of attention. This hypothesis was partially supported. Trial 1 of the HVLT-R correlated highest with the Digit Span subtest ($r = .570$). When compared with the Digits Forward and Digits Backward components of the subtest, Trial 1 of the HVLT-R a significant correlation was found with Digits Forward ($r = .562$) and with Digits Backward ($r = .453$). Although each of these correlation coefficients is significant, they account for only 33% of the shared variance between H-1 and Digit Span and 32% between H-1 and Digits Forward.

Summary and Implications of the Current Findings

The first edition of the HVLT has now been in existence for 17 years and the revised edition for 10 years. It has been well-received by many neuropsychologists as a useful measure of immediate recall, delayed recall, learning, and recognition. Yet, the number of studies focusing on the reliability, validity and utility of this measure has been relatively sparse. The current study was implemented in order to advance the knowledge in the field of neuropsychology concerning the usefulness of the HVLT-R as a measure of verbal memory with individuals who have suffered a mild brain injury.

For standardization data to be useful to the neuropsychologist, the norms must accurately reflect the population with which the norms are intended to be used. If performance on a test is affected by characteristics such as age, education level, race, or gender, yet the standardization data doesn't reflect this, patients can be misclassified for diagnostic purposes. In rehabilitation settings, inaccuracies concerning the patient's progress in treatment can be the result of norms that are not accurate. Of the 5 sets of normative data for the HVLT-R currently in existence, only 2 sets of norms (Brandt & Benedict, 2001; Benedict et al. 1998) provide comparison scores from the late teen years to the late 80s or older. Between these two sets of norms, however, there is a discrepancy in how tight the age ranges need to be set. The current study advances the research with the finding that in a large sample of mild traumatic brain injured individuals, the subjects can be separated into three broad age categories: the young (18-29 years), the middle-aged (30-69 years) and the elderly (70 and older). In this respect, the results are in closer agreement with the Benedict and associates (1998) norms. This is useful information for clinicians to know when choosing which sets of norms more accurately reflect their patients.

The current study also found differences on HVLT-R performance based on level of education. We categorized subjects into three categories: (1) those individuals with less than a high school education, (2) high school graduates, and (3) those with more than high school graduates. We found that, with this very large sample of subjects suffering acute mild traumatic brain injury, two groups emerged: (1) subjects with 12 years or less of formal education, and (2) subjects with more than 12 years of formal education. Although level of education was considered by several sets of HVLT-R normative data, only one set of standardization data (Friedman et al., 2002) actually included an education correction for the standard scores. Friedman and her associates also included a gender correction component as well. Her norms, however, were especially designed to be used to assess African Americans, so this source of norms is limited. Future norming of the HVLT-R should take into consideration the studies, including the current study, that have found level of education to be associated with HVLT-R performance.

Our comparison of the HVLT-R tasks with the Logical Memory tasks with resulted in the HVLT-R placing a substantially higher percentage of subjects in the impaired range on each HVLT-R task compared to the Logical Memory tasks. The lone exception is the percentage retention scores on HVL-R and Logical Memory. A high percentage of subjects scored in the impaired range on

Logical Memory percentage retention. Given the discrepancy between these two measures in concluding that subjects are experiencing impairment in verbal memory functioning, the implications of the current study's results are important for both individual and societal concerns.

From an economic standpoint, the financial cost to society of increased number of patients receiving rehabilitation or perhaps deemed disabled due to scores on the HVLT-R could be significant. The lifetime costs for a head injury are roughly the same despite differences in injury severity (Max, MacKenzie, & Rice, 1991). Max and colleagues looked at data from the year 1985 and found that the average lifetime cost of a mild head injury was approximately $77,000. The average cost of a severe head injury was approximately $93,000. Inability to return to pre-accident employment or any type of employment is a possibility as well.

On the other hand, without the use of the HVLT-R, or similar word-list measures, individuals who need treatment may be overlooked if only prose-story verbal memory tasks are utilized in the assessment of mild-TBI patients. Furthermore, use of only the Logical Memory subtests in the assessment of mild-TBI could result in the denial of Social Security disability benefits for individuals who are actually in need of them. This could put quite a financial burden on an individual and family who is faced with the necessity of returning to work coupled with the loss of productivity.

In a clinical setting such as an acute-care hospital, clinicians may be concerned about the accuracy of paragraph-story memory tasks in the screening for verbal memory deficits in patients suspected of recently suffering a closed head injury. The preference may be to add a verbal memory word-list measure, such as the HVLT-R, to the battery of neuropsychological screening tools. This may alert the clinician to verbal memory deficits that may not be uncovered via paragraph-story memory measures alone. Although the current study revealed that the HVLT-R is much more likely than Logical Memory to identify subjects as impaired, the study cannot confirm that these subjects are indeed impaired. In order to determine whether or not the HVLT-R correctly assessed subjects as impaired or alternatively that the Logical Memory assessed subjects as not impaired, it would be necessary to follow-up with subjects over a period of time subsequent to their release from the hospital. If the subjects identified as impaired by the HVLT-R were unable to resume their pre-accident activities, such as returning to school or returning to their prior employment, and if the subject continued to have difficulty with memory in the months following their accident, this would tend to support the assertion that the HVLT-R is a more sensitive measure of brain impairment. However, such a follow-up study was beyond the scope of the present study and we are, therefore, unable to make absolute assertions about the sensitivity of the HVLT-R to brain injury.

Although new neuropsychological tests are being developed every year, and old tests are often revised, it appears to be a useful project to investigate how current tests can be used to measure constructs it was not originally intended to measure or perhaps to assess individuals that it was not originally normed for. Our attempt to correlate the first trial of the HVLT-R with a focused attention measure, Digit Span Forward, had mixed results. There was a modest correlation between the two tasks, but further research is needed. Logistic regression analyses found that a significant portion of impairment on Trial 1 and Total Recall Score of the HVLT-R is not accounted for by the memory measures used as predictors, including Digits Forward. It should be noted that, if the first trial of the HVLT-R is to be used as a measure of focused attention, the neuropsychologist would have to search for norms because the HVLT-R Professional Manual (Brandt & Benedict, 2001) does not contain normative data for the first trial. Brandt (1991) published norms for the first trial, but he did not find any age or education differences with his sample, so the norms are not stratified by age or education. However, Benedict and Brandt (1998) published norms for the first trial as well as several other scores that are not found in the Professional Manual.

The majority of the research on the HVLT-R and its predecessor, the HVLT, has focused on its use with patients suffering dementia. Research on its utilization with head injured patients is almost non-existent. Based on this study, an argument could be made for the inclusion of the HVLT-R in a standard battery for neuropsychologists to utilize in the screening for verbal memory deficits in patients with acute brain trauma. The substantial number of patients scoring in the impaired range actually confirms the results of Iverson and associates (2000) who found percentages of subjects scoring in the impaired range on the HVLT-R to be very similar to the current study and percentages of subjects scoring in the impaired range on Logical Memory to be very similar to the current study.

However, Iverson and colleagues believed that those results were an anomaly due to the average level of education of their sample falling below the average level of education in the HVLT-R norms. Instead, a pattern may be emerging in the research with mild brain injured patients in which substantially more subjects are doing very poorly on the HVLT-R and not Logical Memory. Although our current study attempted to explore whether this difference was due to the HVLT-R measuring more frontal lobe activity compared to the Logical Memory, our results did not bear this out. Instead, our conclusions point more towards the possibility of the HVLT-R tasks being more sensitive to global brain impairment rather than just the frontal lobes. In this respect, our results support the conclusions of Busch and associates (2005) and not the conclusions of Tremont and colleagues (2000). However, there are other possible interpretations of our findings. There may have been psychological factors that influenced these differences, such as decreased motivation or increased depressive symptomatology. Another possibility is that the HVLT-R may simply be a more difficult task than the Logical Memory test for mild brain injured individuals to master. Our study was not set up to evaluate these possibilities.

Despite these limitations, the sensitivity of the HVLT-R to verbal memory difficulty in mild brain injured individuals provides important support for the argument that both the HVLT-R and the Logical Memory subtest should be included in the neuropsychological battery for the screening of brain injury in the acute care setting. The implications of the current study are that utilizing the Logical Memory subtests as the only verbal memory tools in screening for impairment may be risky because there is a chance that this measure alone may increase the false negatives (patients who actually are experiencing verbal memory impairment that is not identified by the Logical Memory subtests).

Limitations of the Current Study

There are several limitations to the current study. First, the use of archival data makes random assignment impossible. This is a necessary evil, of course, when examining the effects of brain injury since it is not possible to pre-assign subjects to a brain injury group versus a control group.

Second, the study only utilized Form 1 of the HVLT-R even though there are 6 forms altogether. Studies have shown that the six forms are essentially equivalent. However, it is not known whether any of the findings in the current study can be generalized to the alternate forms.

Third, because racial groups other than Caucasians were so highly underrepresented in the current study, it was not possible to determine the extent that race plays on performance of the HVLT-R tasks when an individual suffers a mild brain injury. This limitation is partly due to the limitations of utilizing archival data, although necessary in head injury research.

A fourth and important limitation concerns the limited measures used to create our frontal lobe groups. Ideally, the battery to create those groups would not only consist of the COWAT and Trails B, but also other well-known frontal lobe measures such as the Category Test, Stroop Test, or the Wisconsin Card Sorting Test, as well as PET-scans to document the degree of frontal lobe dysfunction. The current study was limited in not having access to any of these additional measures. Future studies would do well to consider creating frontal lobe groups with more than the results of two tasks.

Directions for Future Research

The current study was not set up to evaluate possible reasons for the discrepancy in performance on the HVLT-R and Logical Memory measures such as the detection of global cognitive impairment, emotional alteration, motivation and drive states, or test difficulty. Future research may focus on these questions.

Future research may want to further explore the difference of conclusions between the Tremont et al. (2000) study, the Busch (2005) and our current study in terms of the frontal lobe involvement in word list versus prose recall verbal memory assessment. We were only able to access two frontal lobe measures (COWAT and Trails B) to create our groups. This is a limitation of the study because executive functioning involves several processes and the COWAT and Trails B alone do not tap most of them. Ideally, future research will create frontal lobe groups to answer this question by utilizing several measures, including neuroimaging measures such as f-MRI and CT-scans, as well as neuropsychological measures such as the Wisconsin Card Sorting Test, Category Test, Trails B, COWAT, and Stroop Color and Word Test, among others.

A few other directions for future research based on the current study's findings are evident. Although the association is inconsistent, there is a growing body of research linking traumatic brain injury to the subsequent onset of Alzheimer’s disease (see Kemp, Goulding, Spencer, & Mitchell, 2005; Lye & Shores, 2000; Jellinger, Paulus, Wrocklage, & Litvan, 2001). Given the relatively high frequency of elderly subjects who scored zero on the delayed recall of the HVLT-R, an interesting future study may focus on the ability of the HVLT-R to predict rapid onset of Alzheimer’s symptomatology in an elderly sample post closed head injury.

Future research may focus on replicating the current findings with the alternate forms of the HVLT-R. Also, exploring the effects of gender and race on HVLT-R performance with individuals suffering from mild brain trauma may be an avenue for future research. Furthermore, more research is needed to determine whether Trial 1 of the HVLT-R could actually be utilized as an attentional measure and, if so, standardization data would need to be established.

REFERENCES

American Psychiatric Association. (1994). *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.). Washington, DC: Author.

Anderson, R.M., Jr. (1994). *Practitioner's guide to clinical neuropsychology*. New York: Plenum Press.

Atkinson, R.C., & Shiffrin, R.M. (1968). Human memory: A proposed system and its control processes. In K.W. Spence & J.T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory* (pp. 89-195). New York: Academic Press.

Baddeley, A. (1986). *Working memory*. Oxford, Clarendon Press.

Baddeley, A. (1994). Working memory: The interface between memory and cognition. In D.L. Schacter & E. Tulving (Eds.), *Memory systems 1994*. Cambridge, MA: The MIT Press.

Baddeley, A. (2000). Short-term and working memory. In E. Tulving & F.I.M. Craik (Eds.), *The Oxford handbook of memory*. Oxford: Oxford University Press.

Baddeley, A.D., & Hitch, G.J. (1974). Working memory. In G. Brown (Ed.), *The Psychology of Learning and Motivation, vol. 8.* (47-90). New York: Academic Press.

Baddeley, A.D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior, 14,* 575-589.

Beery, K.E., & Buktenica, N.A. (1982). *Revised administration, scoring, and teaching manual for the Developmental Test of Visual-Motor Integration.* Cleveland, OH: Modern Curriculum Press.

Benedict, R.H.B. (1997). *The Brief Visuospatial Memory Test-Revised: Professional manual*, Odessa, FL: Psychological Assessment Resources.

Benedict, R.H.B., Schretlen, D., Groninger, L., & Brandt, J. (1998). Hopkins Verbal Learning Test-Revised: Normative data and analysis of inter-form and test-retest reliability. *The Clinical Neuropsychologist, 12*, 43-55.

Bennett, T.L., Dittmar, C., & Ho, M.R. (1997). The neuropsychology of traumatic brain injury. In A.M. Horton, Jr., D. Wedding, & J. Webster (Ed.), *The neuropsychology handbook: Vol. 2: Treatment issues and special populations* (2nd ed., pp. 123-172). New York: Springer Publishing Company.

Benton, A.L., & Hamsher, K. (1989). *Multilingual Aphasia Examination* (2nd ed.). Iowa City, IA: AJA Associates.

Bigler, E.D. (1990). Neuropathology of traumatic brain injury. In Bigler, E.D. (Ed.), *Traumatic brain injury* (pp. 13-49). Austin, TX: Pro-Ed, Inc.

Bond, M.R. (1986). Neurobehavioral sequelae of closed head injury. In I. Grant & K.M. Adams (Eds.), *Neuropsychological assessment of neuropsychiatric disorders* (pp. 347-373). New York: oxford University Press.

Borkowsi, J.G., Benton, A.L., & Spreen, O. (1967). Word fluency and brain damage. *Neuropsychologia, 5*, 135-140.

Brandt, J. (1991). The Hopkins Verbal Learning Test: Development of a new Memory test with six equivalent forms. *The Clinical Neuropsychologist, 5*, 125-142.

Brandt, J., & Benedict, R.H.B. (2001). *Hopkins Verbal Learning Test-Revised*. Lutz, FL: Psychological Assessment Resources.

Bruce, J.M., & Echemendia, R.J. (2003). Delayed-onset deficits in verbal encoding strategies among patients with mild traumatic brain injury. *Neuropsychology, 17*, 622-629).

Busch, R.M., Booth, J.E., McBride, A., Vanderploeg, R.D., Curtiss, G., & Duchnick, J.J. (2005). Role of executive functioning ni verbal and visual memory. *Neuropsychology, 19*, 171-180.

Buschke, H. (1973). Selective reminding for analysis of memory and learning. *Journal of Verbal Learning and Verbal Behavior, 12*, 543-550.

Centers for Disease Control and Prevention: United States Department of Health and Human Services/NATIONAL CENTER FOR INJURY PREVENTION AND CONTROL. (2008). *TBI General Fact Sheet about Traumatic Brain Injury, including information and statistics*. Retrieved June 10, 2008. from http://www.cdc.gov/ncipc/tbi/TBI.htm.

Craik, F.I., & Lockhart, R.S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11,* 671-684.

Crum, R.M., Anthony, J.C., Bassett, S.S., & Folstein, M.F. (1993). Population-based norms for the Mini-Mental State Examination by age and educational level. *Journal of the American Medical Association, 269*, 2386-2391.

Cullum, C.M., Kuck, J., & Ruff, R.M. (1990). Neuropsychological assessment of traumatic brain injury in adults. In E.D. Bigler (Ed.), *Traumatic brain injury: Mechanisms of damage, assessment, intervention, and outcome* (129-163). Austin, TX: Pro-Ed.

Damasio, A.R., Eslinger, P., Damasio, H., Van Hoesen, G.W., & Cornell, S. (1985). Multimodal amnesic syndrome following bilateral temporal and basal forebrain damage. *Archives of Neurology, 42*, 252-259.

Delis, D.C., Kramer, J.H., Kaplan, E., & Ober, B.A. (1987). *California Verbal Learning Test*. New York: The Psychological Corporation.

Ellis, A.W., & Young, A.W. (1988). *Human cognitive neuropsychology*. London: Laurence Erlbaum Associates.

Folstein, M.F., Folstein, S.E., & McHugh, P.R. (1975). Mini-mental state: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12*, 189-198.

Frank, R.M., & Byrne, G.J. (2000). The clinical utility of the Hopkins Verbal Learning Test as a screening test for mild dementia. *International Journal of Geriatric Psychiatry, 15*, 317-324.

Franzen, M.D. (2000). *Reliability and validity in neuropsychological assessment* (2nd ed.). New York: Kluwer Academic/Plenum Publishers.

Franzen, M.D., & Berg, R.A. (1998). *Screening children for brain impairment* (2nd ed.). New York: Springer Publishing Co.

Franzen, M.D., & Wilhelm, K.L. (1998). Conceptual foundations of ecological validity in neuropsychological assessment. In R.J. Sbordone & C.J. Long (Eds.), *Ecological validity of neuropsychological testing* (91-112). Boca Raton, FL: St. Lucie Press.

Friedman, M.A., Schinka, J.A., Mortimer, J.A., & Graves, A.B. (2002). Hopkins Verbal Learning Test-Revised: Norms for elderly African Americans. *Clinical Neuropsychologist, 16*, 356-372.

Gathercole, S.E. (1997). Models of verbal short-term memory. In M.A. Conway, (Ed.), *Cognitive Models of Memory*. Cambridge, Massachusetts: The MIT Press.

Gennarelli, T.A. (1986). Mechanisms and pathophysiology of cerebral concussion. *Journal of Head Trauma Rehabilitation, 1*, 23-29.

Gennarelli, T.A., & Graham, D.I. (2005). Neuropathology. In J.M. Silver, T.W. McAllister, & S.C. Yudofsky (Eds.), *Textbook of Traumatic Brain Injury* (pp. 27-50), Washington, D.C.: American Psychiatric Publishing, Inc.

Goldstein, F.C., & Levin, H.S. (1990). Epidemiology of traumatic brain injury: Incidence, clinical characteristics, and risk factors. In Bigler, E.D. (Ed.), *Traumatic brain injury* (pp. 51-67). Austin, TX: Pro-Ed, Inc.

Goldstein, G. (1998). Functional considerations in neuropsychology. In R.J. Sbordone & C.J. Long (Eds.), *Ecological validity of neuropsychological testing*. Boca Raton, FL: St. Lucie Press.

Goodwin, D.M. (1989). *A dictionary of neuropsychology*. New York: Springer-Verlag.

Goodyear, B., & Umetsu, D. (2002). Selected issues in forensic neuropsychology. In B.V. Dorsten (Ed.), *Forensic psychology: From classroom to courtroom* (pp. 283-313). New York: Kluwer Academic/Plenum Publishers.

Green, B.S., Stevens, K.M., & Wolf, T.D.W. (1997). *Mild traumatic brain injury: A therapy and resource manual*. San Diego: Singular Publishing Group, Inc.

Guilmette, T.J., & Rasile, D. (1995). Sensitivity, specificity, and diagnostic accuracy of three verbal memory measures in the assessment of mild brain injury. *Neuropsychology, 9*, 338-344.

Guskiewicz, K.M., Ross, S.E., & Marshall, S.W. (2001). Postural stability and neuropsychological deficits after concussion in collegiate athletes. *Journal of Athletic Training, 36*, 263-273.

Hanlon, R.T.E., Demery, J.A., Martinovich, Z., & Kelly, J.P. (1999). Effects of acute injury characteristics on neuropsychological status and vocational outcome following mild traumatic brain injury. *Brain Injury, 13*, 873-887.

Hartlage, L.C., Durant-Wilson, D., & Patch, P.C. (2001). Persistent neurobehavioral problems following mild traumatic brain injury. *Archives of Clinical Neuropsychology, 16*, 561-570.

Hogervorst, E., Combrinck, M., Lapuerta, P., Rue, J., Swales, K., & Budge, M. (2002). The Hopkins Verbal Learning Test and screening for dementia. *Dementia and Geriatric Cognitive Disorders, 13*, 13-20.

Iverson, G.L., Lovell, M.R., & Smith, S.S. (2000). Does brief loss of consciousness affect cognitive functioning after mild head injury? *Archives of Clinical Neuropsychology, 15*, 643-648.

Jellinger, K.R., Paulus, W., Wrocklage, C., & Lituan, I. (2001). Traumatic brain injury as a risk factor for Alzheimer's disease. Comparison of two retrospective autopsy cohorts with evaluation of ApoE genotype. *BMC Neurology, 1*, 1-3.

Kaplan, E., Fein, D., Morris, R., & Delis, D. (1991). *WAIS-R-NI Manual: WAIS-R as a neuropsychological instrument*. San Antonio: Psychological Corporation.

Kaplan, E.F., Goodglass, H., & Weintraub, S. (1983). *The Boston Naming Test* (2nd ed). Philadelphia, PA: Lea & Febiger.

Keiska, M.A., Shore, D.L., & Hamilton, J.M. (2007). The role of depression in verbal memory following traumatic brain injury. *The Clinical Neuropsychologist, 21*, 744-761.

Kemp, S., Goulding, P., Spencer, J., & Mitchell, A.J. (2005). Unusually rapid and severe cognitive deterioration after mild traumatic brain injury. *Brain Injury, 19*, 1269-1276.

Kraus, J.F., & Chu, L.D. (2005). Epidemiology. In J.M. Silver, T.W. McAllister, & S.C. Yudofsky (Eds.), *Textbook of Traumatic Brain Injury* (pp. 3-26), Washington, D.C.: American Psychiatric Publishing, Inc.

Kuslansky, G., Katz, M., Verghese, J., Hall, C.B., Lapuerta, P., LaRuffa, G., & Lipton, R.B. (2004). Detecting dementia with the Hopkins Verbal Learning Test and the Mini-Mental State Examination. *Archives of Clinical Neuropsychology, 19*, 89-104.

Lacritz, L.H., & Cullum, C.M. (1998). The Hopkins Verbal Learning Test and CVLT: A preliminary comparison. *Archives of Clinical Neuropsychology, 13*, 623-628.

Levin, H.S., Lilly, M.A., Papanicolaou, A., & Eisenberg, H.M. (1992). Posttraumatic and retrograde amnesia after closed head injury. In Squire, L.R., & Butters, N. (Eds.), *Neuropsychology of Memory* (pp. 290-308). New York: The Guilford Press.

Levin, H.S., O'Donnell, V.M., & Grossman, R.G. (1979). The Galveston Orientation and Amnesia Test: A practical scale to assess cognition after head injury. *Journal of Nervous and Mental Disease, 167*, 675-684.

Lezak, M.D. (1995). *Neuropsychological assessment* (3rd ed.). New York: Oxford University Press.

Long, C.J. (1998). Neuropsychological tests: A look at our past and the impact that ecological issues may have on our future. In R.J. Sbordone & C.J. Long (Eds.), *Ecological validity of neuropsychological testing*. Boca Raton, FL: St. Lucie Press.

Lovell, M.R., Iverson, G.L., Collins, M.W., McKeag, D., & Maroon, J.C. (1999). Does loss of consciousness predict neuropsychological decrements after concussion? *Clinical Journal of Sports Medicine, 9*, 193-198.

Lucas, J.A. (1998). Traumatic brain injury and postconcussive syndrome. In P.J. Snyder & P.D. Nussbaum (Eds.), *Clinical neuropsychology: A pocket handbook for assessment*. Washington, D.C.: American Psychological Association.

Lye, T.C., & Shores, E.A. (2000). Traumatic brain injury as a risk factor for Alzheimer's disease: A review. *Neuropsychology Review, 10*, 115-129.

Manns, J.R., & Squire, L.R. (2002). The medial temporal lobe and memory for facts and events. In A.D. Baddeley, M.D. Kopelman, & B.A. Wilson (Eds.), *The Handbook of Memory Disorders*, 2nd Edition (pp. 81-99). West Sussex, England: John Wiley & Sons Ltd.

Marin, R.S., & Chakravorty, S. (2005). Disorders of diminished motivation. In J.M. Silver, T.W., McAllister, & S.C. Yudofsky (Eds,), *Textbook of traumatic brain injury*. Washington, DC: American Psychiatric Publishing, Inc.

Max, W., MacKenzie, E., & Rice, D. (1991). Head injuries: Cost and consequences. *Journal of Head Trauma Rehabilitation, 6*, 76-91.

Miller, G.A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*, 81-97.

Mitrushina, M., Boone, K.B., Razani, J., & D'Elia, L.F. (2005). *Handbook of normative data for neuropsychological assessment* (2nd ed.). New York: Oxford University Press.

Rains, G.D. (2002). *Principles of human neuropsychology*. Boston: McGraw-Hill Companies, Inc.

Rasmusson, D.X., Bylsma, F.W., & Brandt, J. (1995). Stability of Performance on the Hopkins Verbal Learning Test. *Archives of Clinical Neuropsychology, 10*, 21-26.

Reitan, R.M. (1958). Validity of the Trail Making Test as an indicator of organic brain damage. *Perceptual and Motor Skills, 8*, 271-276.

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

Rimel, R.W. (1981). A prospective study of patients with central nervous system trauma. *Journal of Neurosurgical Nursing, 13*, 132-141.

Rimel, R.W., Giordani, B., Barth, J.T., Boll, T.J., & Jane, J.A. (1981). Disability caused by minor head injury. *Neurosurgery, 9*, 221-228.

Robinson, R.G., & Jorge, R.E. (2005). Mood disorders. In J.M. Silver, T.W. McAllister, & S.C. Yudofsky (Eds.), *Textbook of Traumatic Brain Injury* (pp. 201-212), Washington, D.C.: American Psychiatric Publishing, Inc

Rosen, W.G. (1980). Verbal fluency in aging and dementia. *Journal of Clinical Neuropsychology, 2*, 135-146.

Rutherford, W.H. (1989). Postconcussion symptoms: Relationship to acute neurological indices, individual differences, and circumstances of injury. In H.S. Levin, H.M. Eisenberg, & A.L. Benton (Eds.), *Mild head injury* (pp. 217-228). New York: Oxford University Press.

Sbordone, R.J. (1997). The ecological validity of neuropsychological testing. In A.M. Horton, Jr., D. Wedding, & J. Webster (Eds.), *The neuropsychology handbook volume 1: Foundations and assessment* (2nd ed.). New York: Springer Publishing Company.

Sbordone, R.J. (1998). Ecological validity: Some critical issues for the neuropsychologist. In R.J. Sbordone & C.J. Long (Eds.), *Ecological validity of neuropsychological testing*. Boca Raton, FL: St. Lucie Press.

Sbordone, R.J. & Long, C.J. (1998). Preface. In R.J. Sbordone & C.J. Long (Eds.), *Ecological validity of neuropsychological testing*. Boca Raton, FL: St. Lucie Press.

Selzer, M.E. (1995). Introduction: Epidemiology and pathophysiology of traumatic brain injury. *Journal of Neurology Rehabilitation, 9*, 55-60.

Shapiro, A.M., Benedict, R.H.B., Schretlen, D., & Brandt, J. (1999). Construct and concurrent validity of the Hopkins Verbal Learning Test-Revised. *The Clinical Neuropsychologist, 13*, 348-358.

Sheid, R., Preul, C., Gruber, O., Wiggins, C., & von Cramon, D.Y. (2003). Diffuse axonal injury associated with chronic traumatic brain injury: Evidence from T2*-weighted gradient-echo imaging at 3 T [Electronic version]. *American Journal of Neuroradiology, 24*, 1049-1056.

Shutter, L.A., Jallo, J.I., & Narayan, R.K. (2000). Traumatic brain injury. In R.W. Evans, D.S. Baskin, & F.M. Yatsu (Eds.), *Prognosis of neurological disorders* (2nd ed.) (335-365). New York: Oxford University Press.

Sidman, M., Stoddard, L.T., & Mohr, J.P. (1968). Some additional quantitative observations of immediate memory in a patient with bilateral hippocampal lesions. *Neuropsychologia, 35*, 359-367.

Sosin, D.M., Sniezek, J.E., & Thurman, D.J. (1996). Incidence of mild and moderate brain injury in the United States, 1991. *Brain Injury, 10*, 47-54.

Spreen, O., & Strauss, E. (1998). *A compendium of neuropsychological tests* (2nd ed.). New York: Oxford University Press.

Squire, L.R., & Kandel, E.R. (2008). *Memory: From mind to molecules*. New York: Roberts and Company Publishers.

Strub, R.L., & Black, F.W. (1988). *Neurobehavioral disorders: A clinical approach*. Philadelphia: F.A. Davis Company.

Teasdale, G., & Jennett, B. (1974). Assessment of coma and impaired consciousness: A practical scale. *Lancet, 2*, 81-84.

Tranel, D., & Damasio, A.R. (2002). Neurobiological foundations of human memory. In A.D. Baddeley, M.D. Kopelman, & B.A. Wilson (Eds.), *The Handbook of Memory Disorders*, 2nd Edition (pp. 17-56). West Sussex, England: John Wiley & Sons Ltd.

Tremont, G., Halpert, S., Javorsky, D.J., & Stern, R.A. (2000). Differential impact of executive dysfunction on verbal list learning and story recall. *Clinical Neuropsychologist, 14*, 295-302.

United States Census Bureau. (2007).*Educational attainment in the United States: 2007*. Retrieved June 13, 2008. from http://www.census.gov/population/www/socdemo/education/cps2007.html.

Vallar, G., & Papagno, C. (2002). Neuropsychological impairments of verbal short-term memory. In A.D. Baddeley, M.D., Kopelman, & B.A. Wilson (Eds.), *The handbook of memory disorders* (2nd ed.). West Sussex, England: John Wiley & Sons, Ltd.

Vanderploeg, R.D., Schinka, J.A., Jones, T., Small, B.J., Graves, A.B., & Mortimer, J.A. (2000). Elderly norms for the Hopkins Verbal Learning Test-Revised. *The Clinical Neuropsychologist, 14*, 318-324.

Waugh, N.C., & Norman, D.A. (1965). Primary memory. *Psychological Review, 72*, 89-104.

Wechsler, D. (1981). *Wechsler Adult Intelligence Scale-Revised*. San Antonio: The Psychological Corporation.

Wechsler, D. (1987). *Wechsler Memory Scale-Revised: Manual*. San Antonio: The Psychological Corporation.

Wechsler, D. (1997). *WAIS-III WMS-III technical manual*. San Antonio: The Psychological Corporation.

Williamson, D.J.G., Scott, J.G., & Adams, R.L. (1996). Traumatic brain injury. In R.L. Adams, O.A. Parsons, J.L. Culbertson, & S.J. Nixon, (Eds.), *Neuropsychology for clinical practice: Etiology, assessment, and treatment of common neurological disorders* (pp. 9-64). Washington, DC: American Psychological Association.

Wilson, B., Cockburn, J., & Baddeley, A.D. (1985). *The Rivermead Behavioural Memory Test*. Bury St. Edmunds, Suffolk: Thames Valley Test Company.

Appendix A

Results of Age Group Difference on HVLT-R Performance

Table A1.

Age Group Differences on HVLT-R Total Recall

Age Groups	U	Level of Significance (*p*)	Effect Size (r)
1 versus 2 (n = 217)	4009.50	ns (>.05)	-.07
1 versus 3 (n = 175)	2434.00	<.01	-.20
1 versus 4 (n = 194)	2589.50	=.001	-.24
1 versus 5 (n = 153)	1804.00	=.001	-.27
1 versus 6 (n = 94)	639.50	=.001	-.35
1 versus 7 (n = 105)	304.00	<.001	-.67
1 versus 8 (n = 84)	124.00	<.001	-.70
2 versus 3 (n = 284)	8361.50	<.05	-.13
2 versus 4 (n = 303)	9066.50	<.01	-.18
2 versus 5 (n = 262)	6244.00	<.01	-.19
2 versus 6 (n = 203)	2177.50	=.001	-.23
2 versus 7 (n = 214)	1192.50	<.001	-.53
2 versus 8 (n = 193)	469.00	<.001	-.51
3 versus 4 (n = 261)	8023.50	ns (>.05)	-.05
3 versus 5 (n = 220)	5499.00	ns (>.05)	-.07
3 versus 6 (n = 161)	1937.50	ns (>.05)	-.15
3 versus 7 (n = 172)	1111.00	<.001	-.51

Table A1. (Cont'd)

Age Group Differences on HVLT-R Total Recall

Age Groups	U	Level of Significance (*p*)	Effect Size (r)
3 versus 8 (n = 151)	458.00	<.001	-.52
4 versus 5 (n = 239)	6729.50	ns (>.05)	-.03
4 versus 6 (n = 180)	2393.00	ns (>.05)	-.11
4 versus 7 (n = 191)	1426.00	<.001	-.46
4 versus 8 (n = 170)	632.50	<.001	-.46
5 versus 6 (n = 139)	1760.00	ns (>.05)	-.09
5 versus 7 (n = 150)	1126.50	<.001	-.45
5 versus 8 (n = 129)	507.50	<.001	-.48
6 versus 7 (n = 91)	548.00	<.001	-.40
6 versus 8 (n = 70)	245.00	<.001	-.50
7 versus 8 (n = 81)	679.50	ns (>.05)	-.09

Note: Group 1 = 18-19 years; Group 2 = 20-29 years; Group 3 = 30-39 years; Group 4 = 40-49 years; Group 5 = 50-59 years; Group 6 = 60-69 years; Group 7 = 70-79 years; Group 8 = 80+ years.

Table A2.

Age Group Differences on HVLT-R Delayed Recall

Age Groups	U	Level of Significance (*p*)	Effect Size (r)
1 versus 2 (n = 217)	3969.50	ns (>.05)	-.07
1 versus 3 (n = 175)	2542.00	<.05	-.18
1 versus 4 (n = 194)	2723.00	<.01	-.22
1 versus 5 (n = 153)	1752.50	=.001	-.29
1 versus 6 (n = 94)	727.00	<.01	-.28
1 versus 7 (n = 105)	574.00	<.001	-.51
1 versus 8 (n = 84)	293.00	<.001	-.53
2 versus 3 (n = 284)	8845.00	ns (>.05)	-.09
2 versus 4 (n = 303)	9622.00	<.05	-.14
2 versus 5 (n = 262)	6181.50	=.001	-.20
2 versus 6 (n = 203)	2543.50	<.05	-.15
2 versus 7 (n = 214)	2085.00	<.001	-.37
2 versus 8 (n = 193)	1098.50	<.001	-.35
3 versus 4 (n = 261)	7983.50	ns (>.05)	-.05
3 versus 5 (n = 220)	5142.50	ns (>.05)	-.12
3 versus 6 (n = 161)	2095.50	ns (>.05)	-.10
3 versus 7 (n = 172)	1694.50	<.001	-.36
3 versus 8 (n = 151)	903.50	<.001	-.35
4 versus 5 (n = 239)	6331.50	ns (>.05)	-.07

Table A2. (Cont'd)

Age Group Differences on HVLT-R Delayed Recall

Age Groups	U	Level of Significance (*p*)	Effect Size (r)
4 versus 6 (n = 180)	2588.00	ns (>.05)	-.06
4 versus 7 (n = 191)	1426.00	<.001	-.46
4 versus 8 (n = 170)	1165.00	<.001	-.30
5 versus 6 (n = 139)	1977.00	ns (>.05)	-.00
5 versus 7 (n = 150)	1724.50	=.001	-.26
5 versus 8 (n = 129)	944.00	<.01	-.27
6 versus 7 (n = 91)	727.50	<.05	-.25
6 versus 8 (n = 70)	394.00	<.05	-.30
7 versus 8 (n = 81)	737.00	ns (>.05)	-.03

Note: Group 1 = 18-19 years; Group 2 = 20-29 years; Group 3 = 30-39 years; Group 4 = 40-49 years; Group 5 = 50-59 years; Group 6 = 60-69 years; Group 7 = 70-79 years; Group 8 = 80+ years.

Table A3.

Age Group Differences on HVLT-R Percentage Retention

Age Groups	U	Level of Significance (*p*)	Effect Size (r)
1 versus 2 (n = 217)	3808.00	ns (>.05)	-.10
1 versus 3 (n = 175)	2612.50	ns (>.05)	-.16
1 versus 4 (n = 194)	2875.00	=.01	-.19
1 versus 5 (n = 153)	1808.00	=.001	-.27
1 versus 6 (n = 94)	784.00	<.05	-.23
1 versus 7 (n = 105)	762.50	<.001	-.39
1 versus 8 (n = 84)	417.50	<.001	-.40
2 versus 3 (n = 284)	9329.00	ns (>.05)	-.05
2 versus 4 (n = 303)	10220.50	ns (>.05)	-.09
2 versus 5 (n = 262)	6513.50	<.01	-.16
2 versus 6 (n = 203)	2762.00	ns (>.05)	-.11
2 versus 7 (n = 214)	2697.00	<.001	-.26
2 versus 8 (n = 193)	1517.00	=.001	-.24
3 versus 4 (n = 261)	8028.00	ns (>.05)	-.05
3 versus 5 (n = 220)	5151.00	ns (>.05)	-.12
3 versus 6 (n = 161)	2183.00	ns (>.05)	-.07
3 versus 7 (n = 172)	2076.50	=.001	-.26
3 versus 8 (n = 151)	1180.50	<.01	-.24
4 versus 5 (n = 239)	6348.00	ns (>.05)	-.07

Table A3. (Cont'd)

Age Group Differences on HVLT-R Percentage Retention

Age Groups	U	Level of Significance (*p*)	Effect Size (r)
4 versus 6 (n = 180)	2655.00	ns (>.05)	-.04
4 versus 7 (n = 191)	2650.50	<.01	-.20
4 versus 8 (n = 170)	1483.00	<.05	-.20
5 versus 6 (n = 139)	1938.00	ns (>.05)	-.02
5 versus 7 (n = 150)	2029.00	<.05	-.16
5 versus 8 (n = 129)	1143.50	ns (>.05)	-.17
6 versus 7 (n = 91)	831.50	ns (>.05)	-.16
6 versus 8 (n = 70)	465.00	ns (>.05)	-.19
7 versus 8 (n = 81)	735.00	ns (>.05)	-.03

Note: Group 1 = 18-19 years; Group 2 = 20-29 years; Group 3 = 30-39 years; Group 4 = 40-49 years; Group 5 = 50-59 years; Group 6 = 60-69 years; Group 7 = 70-79 years; Group 8 = 80+ years.

Appendix B

Results of Level of Education Difference on HVLT-R Performance

Table B1.

Education Level Differences on HVLT-R Total Recall

Education Groups	U	Level of Significance (*p*)	Effect Size (r)
LHS versus HS (n = 463)	22044.00	ns (>.05)	-.08
LHS versus MHS (n = 397)	12844.00	<.001	-.28
HS versus MHS (n = 536)	27119.50	<.001	-.20

Table B2.

Education Level Differences on HVLT-R Delayed Recall

Education Groups	U	Level of Significance (*p*)	Effect Size (r)
LHS versus HS (n = 463)	23980.00	ns (>.05)	-.01
LHS versus MHS (n = 397)	14383.50	<.001	-.21
HS versus MHS (n = 536)	27359.00	<.001	-.20

Table B3.

Education Level Differences on HVLT-R Percentage Retention

Education Groups	U	Level of Significance (*p*)	Effect Size (r)
LHS versus HS (n = 463)	24032.00	ns (>.05)	-.01
LHS versus MHS (n = 397)	15752.50	<.05	-.15
HS versus MHS (n = 536)	28522.00	<.001	-.17