WECHSLER MEMORY SCALE RUSSELL REVISION: DIFFERENTIATION OF CLOSED HEAD INJURY, STROKE, DEMENTIA, AND PSEUDODEMENTIA

BY

MARY ANN SCOTT

Bachelor of Arts
University of Missouri - Columbia
Columbia, Missouri
1986

Master of Science
Oklahoma State University
Stillwater, Oklahoma
1988

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the degree of
DOCTOR OF PHILOSOPHY
July, 1992

WECHSLER MEMORY SCALE RUSSELL REVISION: DIFFERENTIATION OF CLOSED HEAD INJURY, STROKE, DEMENTIA, AND PSEUDODEMENTIA

Thesis Approved:

Thesis Advisor

Dean of Graduate College

ACKNOWLEDGEMENTS

I would like to first thank Kenneth Sandvold, Ph.D. without whom this dissertation would not have been possible. I would also like to express my deeply felt appreciation to Russell E. Adams, Ph.D. for generously allowing me to use his archival data and for setting a research example to strive for. Additionally, I would like to thank David Thomas, Ph.D. for his guidance and thoughtful critique of this research. Thanks also to committee members Robert, Webber, Ph.D. and Brent Snow, Ph.D..

I must extend my gratitude to Scott Pitts, computer programer without whose patience and willingness to stay late I would still be sitting in the computer lab trying to analyze my data. Thanks also goes to Andy Cucchiara, SAS expert for his consultation and trouble shooting.

To my parents, Jim and Velma Scott, for instilling in me a strong work ethic; and to my grandmother, Mary Scott, for her support and encouragement. I would also like to thank the members of my graduating class for sharing in this experience. My thanks to my dear friends Debbie Wiggins, Shannon Milner, and Donnel Crow for being there when I needed support. Finally, my heart felt appreciation to my husband, Mehdi Sarvestani, for sticking by me and believing in me. For all of his evenings spent typing and xeroxing my

dissertation and for all the times he picked up the pieces and gave me the courage I needed to continue.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

INTRODUCTION

Over the years the medical and psychological communities have begun to turn to neuropsychologists to tease out the etiology of complex groups of symptoms suggestive of neurological impairment. Neuropsychological tests have gained increasing popularity and respect for their sensitivity to organic processes and their diagnostic accuracy (Heaton, Baade, and Johnson; 1978). As the questions that neuropsychologists are asked to address become increasingly complicated, the proceedures used to make these decisions must become more sophisticated and precise.

Memory has long been recognized as a complex higher cortical function that is very sensitive to both organic and functional processes. Deficits in memory have long been used as indicators of a significant organic or functional process within the context of a mental status exam. Neuropsychological test batteries also typically include an assessment of memory functioning.

Although the Weschler Memory Scale (WMS) has been criticized for its norms and construction, it continues to be the most widely administered test of its kind. Many neuropsychologists have adopted the modifications suggested by Russell in 1975. A newly revised version of the WMS is now available. However, it is suggested by D'Elia, Satz, and Schretlen (1989) that many neuropsychologists will continue to use the original form and simply employ the modifications suggested by Russell (1975).

The WMS Russell revision assesses memory on several different levels and ultimately results in a Memory Quotient (MQ). Although the MQ may be useful in describing the significance of an individual's memory impairment, it is not helpful in describing the nature of the impairment. The present study uses existing WMS profiles to determine whether or not an identifiable pattern of deficit can be linked to a specific eitiology and thus be used to aid in the diagnosis of organic and functional processes. Neuropsychologists are often asked to assess whether or not neurological impairment is the result of a diffuse closed head injury (CHI), a cerebral vascular accident, a dementing process, or is of functional origin. This study attempts to derive a formula based on individual memory functions as measured by the WMS Russell revision that can correctly predict group membership in the aforementioned categories. Such a formula could be used to assist neuropsychologists in the interpretation of a WMS profile.

CHAPTER II

LITERATURE REVIEW

Psychologists have been assessing memory in psychiatric patients since the beginning of this century. However, until Weschler's article was written in 1945 no formal test of memory had ever been fully developed. Weschler felt a standardized memory test was badly needed. Although Wells had put forth a memory test in 1923, it was not fully developed and there had been no attempt to standardize it. Weschler (1945) expanded on Wells' idea of a memory quotient and developed his test over a ten year period of experimenting with measures of memory.

In the early 1900's, memory was poorly understood. Ebbinghaus (1913) pioneered the early work on memory theory with his development of nonsense syllables. He was the first to recognize that prior learning has an indeterminate effect on memory which he controlled for by developing meaningless syllables for use in measuring memory. Ebbinghaus also developed the idea that memory could be measured in terms of "savings". A "savings score" was derived by comparing the number of trials required to learn a novel list of nonsense syllables to the number of trials required to relearn the same list at a later date.

Although his work revolutionized the field of memory research, Ebbinghaus (1913) did not provide a definition of memory and memory theory remained vague. Memory was believed to be a unitary function through which information was cataloged and stored. The content of the information to be remembered was the only factor discussed as having an important effect on memory (Ebbinghaus, 1913). Although Bartlett (1932) provided evidence that memory was more complex than Ebbinghaus had suggested by demonstrating that each individual's memory for a novel story was different, his view on memory did not take hold until the early 1960's.

In Wechsler's (1945) work on the development of a memory measure reference was not made to any theoretical papers of the time. He approached the problem through trial and error. A number of different memory tests were tried. He relied on his own clinical experience with memory measurement and chose the subtests he observed to be the most useful in distinguishing between organics and normals and in diagnosing specific organic processes. Wechsler did not link his test to a theory of memory existing at that time.

The Weschler Memory Scale (WMS) resulted from this work and consists of seven different memory scales: Personal and Current Information (Inf), Orientation (Or), Mental Control (MC), Logical Memory (LM), Memory Span (MS), Visual Reproduction (VR), and Associate Learning (AL). Each scale was included as a means of detecting specific organic populations. The Personal and Current Information scale consists of six questions. Two of the questions ask the subject to provide demographic information about themselves while the other four questions ask them to name federal, state, and local government officials. The orientation scale consists of five items that assess the subjects orientation to time and place. These two scales were included for use with patients that have aphasia or suffer from dementia. Both scales were found to be poor discriminators between normal and near normal groups. The Mental Control Subtest requires the subject to perform three simple mental operations: counting backwards from 20 to one, reciting the alphabet, and counting by threes. Wechsler found this subtest to be valuable in diagnosing subjects in the early stages of an organic brain disease. The Logical Memory subtest consists of two separate passages which are read to the subject. Upon completion of each story the examine is asked to recall as much as possible. This subtest was designed to assess the immediate recall of logical information. The Memory Span subtest was taken directly from the digit span subtest on the Wechsler Bellevue Intelligence Scale. This subtest requires the subject to recall number series of increasing length forward in the first phase of the task and backward in the second phase. Wechsler did not provide a rationale for inclusion of this specific subtest. The Visual Reproduction subtest is comprised of four simple geometric figures. The examine is allowed to look at each figure for ten seconds and is then asked to reproduce it. The figures are scored according to a set of criteria outlined by Wechsler. The rationale for including this subtest was again not provided. The last subtest, Associate Learning is comprised of a series of ten paired associates. Five of the pairs are considered easy because of their common associations and the other five are considered hard as the associations are more remote. The subject is read the list of pairs and told that they will be later asked to provide the associate. The subjects are told whether or not their response is correct and are provided with the correct associate when it is missed. The procedure is repeated with varied orders of the pairs for a total of three presentations. The Associate Learning subtest was included because it was found useful in assessing the retention deficits of Korsakoff's patients.

The WMS was normed on 200 men and women ages 25-50, who were not hospital patients. The Weschler Bellevue Intelligence scale scores were available for 100 of the subjects. Their scores on the performance subtests were found to parallel performance on the WMS. The subjects were grouped into age increments of five years beginning at age 20 and ending at age 64. A constant that would equate the scores on the two tests was empirically derived for each age group and became part of an equation that was used to compute the memory quotient. Therefore, the MQ allows for the variation of the scores with age and equates the performance of each age group on the WMS with their intelligence quotient as measured by the Wechsler Bellvue Intelligence test. Wechsler described the strengths of the instrument by stating that it is brief and easy to administer, has adequate standardization, accounts for the differences in memory due to age, and is comparable to an intelligence quotient. It was Weschler's hope that the WMS would prove useful in differentiating between organic and nonorganic groups (Weschler, 1945).

In the years that followed the development of the WMS, relatively little research was conducted on the scale. However, researchers exploring the nature of intelligence began to study the pattern of performance exhibited by different diagnostic groups on the Wechsler-Bellevue Intelligence Scale (Wechsler,1944; Rappaport, 1945). The results indicated that within a given diagnostic category such as a neuropsychiatric group, patients exhibited similar patterns of performance. These studies suggested that the interrelationships of subtest scores yield valuable diagnostic information. This idea of pattern analysis was later applied to the WMS (Cohen, 1950). Cohen tested his theory that the differences exhibited on subtest scores between established diagnostic groups could aid in diagnosis and advance the understanding of the disorders. The study divided subjects into three diagnostic categories: organics, psychoneurotics, and schizophrenics. The subtest scores for each group were compared through separate analyses of variance. The results indicated that no significant differences in performance existed between the groups on any of the subtests given. Cohen (1950) concludes that the WMS subtest and composite scores do not differentiate between these three diagnostic groups.

However, concluding from this finding that the WMS is not useful in distinguishing between groups is premature because Cohen grouped the subjects into general diagnostic categories. By today's standards, such broad diagnostic groups are viewed as essentially meaningless. Cohen also failed to outline the criteria used for making the diagnoses and did not control for the duration and severity of the disorders. Additionally, he neglected to consider whether or not the subjects were taking psychoactive medications. Therefore, any conclusion that the WMS is not useful in differentiating amoung diagnostic groups is premature.

Parker (1957) included the WMS in an investigation of the diagnostic acuity of various tests used in evaluating brain-injured patients. His experimental group contained 30 male patients that had been diagnosed with brain injuries which had occurred within ten months of the time of testing. The experimental group was compared to a hospitalized control group that was comprised of patients receiving care for physical problems. The two groups were matched for age and IQ. They were then compared with respect to their MQ's and their performance on the visual reproduction subtest. No significant differences were found between the two groups on these two measures. Parker (1957) concluded that the MQ's yielded by the WMS were not valid estimates of the subjects' memory abilities. However, although Parker addressed past problems by including a hospitalized control group, he did not describe the types of physical problems included in the group. He also failed to consider the effect that pain and pain medication might have on memory. Additionally, he failed to consider the severity and time since injury when constructing the brain-injured group. In light of the weaknesses of the study, it is premature to conclude that the WMS MQ is an invalid estimate of memory functioning until these methodological problems have been addressed.

Shontz (1957) addressed the limitations of the WMS for use in evaluating severely impaired patients. He was interested in the extent to which physically handicapping conditions interfere with a patient's ability to demonstrate his or her intellectual abilities. The population addressed in this study were hemiplegics. Shontz (1957) was particularly interested in the problems presented when testing a right hemisphere hemiplegic that suffered from a loss of functioning on the dominant side of his body along with impairments in verbal expression. Fifty hemiplegic patients (17 right hemiplegics and 33 left hemiplegics) were administered the nonverbal, nonmotor Columbia Maturity Scale (CMS) and the Knox Cube Test (KCT) which asked the subject to repeat a tapping pattern which required gross motor movements. Scores for the LM, MC, DSp, and AL subtests on the WMS were available for 31 of the subjects. No significant differences were found between right and left hemisphere hemiplegics on the CMS or the KCT. However, right hemiplegics (left hemisphere lesions) performed significantly poorer on the combined subtests of the WMS. Shontz (1957) concluded that the WMS was not appropriate for use in evaluating hemiplegics because of its high reliance on verbal skills. However, in light of the view of memory as a unitary function that existed in the 1950's, he failed to consider the possibility that left versus right hemisphere injuries might affect memory differentially. The argument that in such cases it is difficult to know whether or not one is measuring a deficit in memory or verbal expression, is an important one and is consistent with the results.

Shortly after Cohen's (1950) article was published, Howard (1950) conducted a study to investigate the utility of the WMS in differentiating between the organically and functionally impaired. The study employed three groups: subjects with both an organic and psychotic diagnosis, subjects with a diagnosis of only psychosis, and an experimental group composed of paretics, encephalitics, and epileptics. A control group of paretics and normals matched for age, length of hospitalization, education, and occupation was also included in the study. The method used in making these diagnoses was not outlined. However, this was the first study to take into account the effect that the length of hospitalization might have on the results. Between group comparisons of the MQ and other subtest scores were made. The study also included the analyses of single items and subcomponents of the subtest scores which included: counting backwards, alphabet, count by three's, LM part A, LM part B, VR for each figure, and AL easy and hard. Theses variables were included to determine whether valuable information was being lost in grouping items together into subtest and looking only at those subtest scores. The study showed that only the paretic group could be differentiated from the other groups on the basis of their WMS subtest scores. Howard (1950) concluded that the WMS may be useful in discriminating individuals with gross brain-damage from matched groups. However, it seems that in cases of gross brain-damage the origin is more apparent, and the WMS is not likely to be used as a diagnostic tool. Count by Three's, AL easy and hard, and VR for Figures B and C1 all reached the .05 level of significance; however, the overall subtests scores were also significant. Howard (1954) followed up on the original study by repeating it with 25 subjects that were given an organic and psychotic diagnosis, and 25 subjects that were diagnosed with psychosis. In this study the subjects were matched for age and length of hospitalization. Additionally, it was reported that most of the subjects had negative results on a recent neurological examination. The results differed from those obtained in the original study, suggesting that the WMS was useful in differentiating between the groups. The results of the study found the following criteria set of criteria useful in differentiating between the groups: MS < 8, Who is the Governor of the patients state of residence (an item from Inf) incorrect, MQ < 75, Inf < 5, VR < 4, VR on figure C-1 scored 0. However, it is important to note that further replications which yielded the same set of criterion variables would have to be produced before these criteria can be utilized.

By the late 1960's and early 1970's memory theory had changed dramatically. Bartlett's (1932) view of memory as a constructive process began to replace the earlier assumptions that memory is a unitary process. As this new view of memory was adopted, memory research became increasingly detailed and specific. Guilford (1967) provided conclusive evidence for treating retention and retrieval as separate memory processes. He also provided research that supported a distinction between semantic memory (memory for ideas) and symbolic memory (memory for words, letters, numbers, etc.)

In a series of research studies, it was demonstrated that memory functions were lateralized to different hemispheres of the brain (Kimura, 1963; Milner, 1968, 1970). Kimura conducted a study on epileptics, most of whom had received a unilateral temporal lobectomy, to investigate the visual deficits of the right hemisphere temporal lobectomy group. The results of the study suggested that the right hemisphere is important for the perception of novel visual information while the left hemisphere may be more important in recognizing familiar visual information. This research study was important as it demonstrated the complexity of memory functions as they exist in the brain. Milner (1968) gave three separate facial recognition tasks to patients with left and right temporal lobectomies and another group of patients with frontal or parietal lobe excisions. For each experiment identical arrays of faces for the learning and recognition trials were employed. The individual subjects and the delay before recognition were the variables that differed beween experiments. The results supported Milner's conclusion that right temporal lobe excision impairs memory for novel faces while left temporal lobectomy does not.

In a later review of the research conducted on patients with temporal lobe lesions, Milner (1970) discussed the differential effects of unilateral versus bilateral temporal lobectomies on memory. Bilateral temporal lobectomies were shown to result in severe impairment in generalized memory functioning. Unilateral temporal lobectomies were found to result in memory impairment limited to the registration of new information for long-term recall.

Black (1973) gave further support to the idea that different types of brain damage produce variable deficits and that brain damage is not a unitary process. In this study, he compared the performances of 50 patients with missile wounds and 50 patients closed head injuries (CHI's) on the WAIS-R, WMS, and various other tests. The results found the CHI group to have significantly lower IQ's and MQ's than the missile wound group. Black concluded that missile wounds do not always cause significant cognitive or memory impairment, and his study offers evidence that different brain injuries result in different residual impairments.

In 1971, Dujovne and Levy proposed that the WMS is multidimensional and conducted a factor analysis on it. Two groups were employed. The "normal" group consisted of 276 individuals with no history of brain injury or psychiatric illness and was made up of 60% females and 40% males. The patient group consisted of 81 individuals with varied diagnoses. The majority of the subjects in this group were described as having acute and chronic brain disorders. They were also dually diagnosed with psychotic, psychoneurotic, and transient situational personality disorders. The factor structure of the normal and patient groups were found to be different. The conclusion that the WMS is multidimensional is consistent with the results. However, this study utilized mixed groups of patients and did not attempt to control for the large differences between the two groups with respect to age, years of education, sex, or length of hospitalization, and the results may have been partly due to these methodological problems. The suggestion that resulting factors should replace the MQ as a method of scoring was premature. The stability of the factors must first be demonstrated through further study before they can be adopted as a scoring method.

A later study contradicted the finding of Dujovne and Levy with results that showed the factor structure of the WMS to be almost identical for mixed psychiatric and organic groups (Kear-Collwell, 1973). This study employed more stringent criteria for the inclusion of subjects in the organic groups and compared the performances of head injured, demented, and a collective group of patients with a severe neurological impairment on the WAIS and WMS with a group of patients referred for testing but for which no organic pathology could be confirmed. A factor analysis revealed almost identical factor structures for the two groups. The study then calculated factor scores for each subject and compared the two groups on the three resulting WMS factors Full scale IQ (FSIQ), Verbal IQ (VIQ), Performance IQ (PIQ), their verbal - performance discrepancy score, and age. Significant differences were found between the two groups on the factors that involved LM, VR, and AL and the factor that was primarily composed of Or and Inf.

The researchers went on to look for differences between specific organic groups. The study found that the mixed neurologically impaired group did not significantly differ from the nonconfirmed group on any of the variables. The CHI group was found to evidence only slight intellectual impairment and showed no significant impairments in memory functioning when compared to the nonconfirmed group. Kear-Collwell (1973) presented a fair discussion of the results which stressed the need to take into account the severity and time since injury when looking at a CHI group. The study concluded that both the localization and extent of neurological impairment must be considered when trying to determine the relationship between organic impairment and cognitive functioning. It was further suggested that the WMS be reorganized into a factor structure and implied that the utility of the MQ is quite limited.

Bachrach and Mintz (1974) compared psychiatric groups with a detectable cerebral dysfunction and those without.

Although the groups were highly heterogeneous, they were carefully matched for psychiatric diagnosis. Bachrach and Mintz (1974) found that four subtests significantly discriminated between the groups: Inf, LM, VR, and AL. It was also observed that VR alone was able to discriminate between impaired and unimpaired subjects with 89% accuracy. The authors failed to offer an explanation for this last finding. The study concluded that the WMS could be used to detect mild cerebral dysfunction in matched groups.

Kijajic (1975) criticized Bachrach and Mintz (1974) for failing to control for age and education and attempted to address the problem of age and education effects by developing equations based on hold and don't hold subtests of the WMS. Two regression equations that could be used to discriminate between the subjects with a statistically significant degree of accuracy resulted. This study appears methodologically sound with the exception of using nonspecific organic groups. The author suggests that these equations be used as a screening device for organic impairment. Although the equations look promising, further research is needed before they can be determined reliable.

Kessner (1973) presented an elaborate information processing theory of memory. The theory separated memory into three components; cue access, short-term memory and long-term memory storage, and a retrieval system. Kessner also identified a number of factors believed to influence memory, such as selective attention, rehearsal, arousal, and consolidation. The paper ties the theory to underlying neural processes by interrelating biological data available from evoked potential and animal research studies with the models presented.

Kessner's (1973) view of memory functioning is radically different from the assumption that memory is a unitary process which was held into the 1960's. With the publication of this article and a change in the underlying assumptions that research was based on, the methodology of previous studies was called into question. The lack of understanding of memory functioning and brain damage is reflected in the methodology of clinical memory studies up until the late 1970's. The groups employed by these studies would be considered essentially meaningless by today's standards. Therefore, the meaning of the results reported in the early studies is difficult to interperet. Additionally, as the design of the WMS was based on false assumptions, its usefulness in evaluating memory impairment was also called into question.

Russell (1975) proposed that the advances made toward the understanding of memory processes and the effects of brain damage demonstrate a need for more precise measures of memory functioning than the original WMS can provide. He suggested that a lateralization component be added to the WMS that would reflect figural versus verbal differences. Russell also proposed that immediate and long-term memory processes should be measured separately. His final suggestion was that if the WMS was to be used in evaluating brain damage, a method of relating it to other measures of brain impairment should be derived.

Russell (1975) introduced a modified administration and scoring procedure that was designed to increase the utility of the WMS and incorporate the components of memory theory discussed above. The method included a 30 minute recall of the logical and visual memory materials, and the score on the second recall could then be divided into the score on the first recall to yield a measure of the percentage of information retained for both types of memory. Russell (1975) chose the logical memory and visual reproduction subtests for this delayed procedure because they are highly affected by brain damage and represent lateralized functions. Russell's procedure also introduced an average impairment rating (RWMS) for the WMS. The procedure was validated on four experimental groups comprised of brain damaged subjects, right-hemisphere damaged subjects, left-hemisphere damaged subjects, and controls. There were no significant differences among the groups on age and education. The subjects were all male. The reliability of the scores was assessed with an internal consistency method. The only score which proved to be of questionable reliability was the figural percent retained score. This finding was believed to result from the decreased sensitivity of this measure when assigning a percent retained score to individuals who had no memory for the figures on either recall. According to the formula used in calculating their percent retained score, they would receive a score of 100%. The validity of using the modified procedure as a measure of organicity was also evaluated. The delayed recall scores for logical (semantic) and visual (figural) memory and the percent retained scores for both logical and visual memory were found to be impaired in brain damaged subjects. However, the percent retained scores were again found to be less sensitive due to the problem of assigning a percent to low scores.

The test's sensitivity to specific brain impairments was assessed by correlating the amount of impairment in these memory tests with the Average Impairment score derived from the Halstead-Reitan battery. The results indicated that the amount of memory impairment is strongly related to the Average Impairment scale produced by brain damage. The correlations were all found to be significant at the .01 level. The Average Impairment and WMS subtest scores were then transformed into comparable Z scores. The scale scores that had been set for the Average Impairment scores were then applied to the memory scores (Russell, 1975). The concept of an impairment rating was first introduced by Russell, Neuringer, and Goldstein (1970). The ability of this new procedure to lateralize memory impairment was also evaluated and found to provide a reliable and valid means of differentiating right from left hemisphere damage. In the discussion it is pointed out that the WMS primarily assesses memory functions located in the temporal lobe. This may be in part a result of Russell's (1975) failure to incorporate a means of distinguishing between storage and retrieval deficits. Although this study provides solutions to many of the problems with the construction of the original WMS, it still ignores an important aspect of more recent memory theory.

In reviews of the WMS literature, several limitations of the WMS were cited (Erickson & Scott, 1977; Prigatano, 1978). Erickson and Scott (1977) stated that the research in the area of neuropsychology that attempts to relate memory functioning to brain functioning continues to raise questions but fails to find conclusive answers. The results of studies conducted in the 1950's and 1960's were cited in discussing the limitations of the WMS and other tests of neuropsychological functioning. However, the methodological problems of these studies were not taken into consideration, The review criticized the WMS for its inability to discriminate between organic and nonorganic patients, poor internal consistency, and its close relationship to measures of intelligence. It was concluded that a new measure of memory based on memory theory of the time was needed. This conclusion is valid; however, the utility of the WMS can not be determined from poorly done studies. Additionally, the review came too early to include any studies that had employed the modified procedure introduced by Russell (1975).

A later review by Prigatano (1978) focused exclusively on the WMS. The WMS was criticized for its lack of representative normative data, its failure to provide standard scores for the subtests, and a lack of studies addressing the distribution of full scale IQ minus MQ scores. Although Prigatano (1978) can also be criticized for failing to discuss the methodological problems of the studies cited, the conclusions were drawn by looking for consistencies across studies. Additionally the review included some studies that had incorporated Russell's delayed recall procedure which was relatively new at the time of the review. This review also listed the strengths of the WMS which included : a) a relatively constant factor structure across studies; b) an observed decline in performance with age as would be expected and; c) experimental support for the WMS as a test of short-term verbal memory. The overall conclusion of the review was that the WMS is valuable and that Wechsler succeeded in picking important subtests for the evaluation of memory despite his lack of knowledge of underlying memory processes. Finally, Prigatano (1978) suggested that the test be reconstructed to reflect the memory theory of the time.

The construct validity of a number of memory testing procedures including selected subtests of the WMS Russell modification (1975) were more recently reviewed (Larrabee, Kane, Shuck, and Francis, 1985). The subtests of the WMS included in the study were LM, VR, and PA. The delayed recall scores for LM and VR were also included in the study. A factor analysis revealed four factors composed of: a) verbal and visual learning and memory; b) visual-perceptual-motor abilities; c) verbal abilities and; d) attention/concentration. Larrabee et al. (1985) concluded that the delayed recall scores of LM and VR were better able to differentiate right from left hemisphere damage than the initial recall of these materials and were more effective discriminators than the Benton Visual Retention Test. The study also found evidence of construct validity for LM and PA. The researchers failed to describe the exact procedure used in administering the delayed recall and did not describe the method used in scoring the LM stories. As Russell (1975) did not provide standard instructions for use in implementing his modifications, this information is important if a study is to be replicated.

Since its development, the WMS has been criticized for failing to provide a comprehensive and representative set of norms. The original norms were based on individuals between the ages of 25 and 50 (Weschler, 1945). The original norms were meant to be provisional, but as late as the 1970's, no new normative studies had been conducted for this young adult population (Prigatano, 1978). Some initial norms for younger groups of adolescents and children were developed by Ivinskis, Adams, and Shaw (1971). Hulica (1966) and Cauthen (1977) attempted to extend the norms to individuals age 60 and over. Hulica (1966) gave the WMS to 237 individuals in the following age groups: 15-17, 30-39, 60-89. The number of subjects in each group was not provided. The results found that the overall performance of younger subjects was superior to that of the older groups as expected. The subtests that contributed the most to the decline in performance were found to be LM, MS, and AL. The subjects were matched on an estimated IQ assessed with the vocabulary subtest of the WAIS; however they were not matched for years of education which remains constant and does not change with age in a given individual.

Cauthen (1977) conducted a similar study on a smaller sample of elderly individuals that employed only 10-15 subjects per age interval. This study also failed to control for education but broke down the normative data by IQ scores. The results showed the performance on VR and the delayed recall of both LM passages to significantly differ between the groups. These three subtests were found to be the most related to age. These studies (Cauthen, 1977; Hulica, 1966; and Ivinskis et. al., 1971) were important but failed to provide age corrections for use at arriving at an MQ for these age groups. In a review of the literature, Prigatano (1978) described these studies as providing little more than guidelines for use with adolescent and elderly populations.

In the 1980's the number of normative studies that were conducted on the WMS had increased. A study of the performance of a healthy elderly population reported norms for this group and addressed the question of whether memory decline is a direct result of aging or is seen only in elderly individuals (age 65 and older) with serious medical conditions (Haaland, Lin, Hunt, and Goodwin, 1983). Only a portion of the WMS, LM and VR with delayed recall were administered. The results showed that performance on tasks measuring immediate and delayed recall of both semantic and figural information significantly declines with age. No significant differences on the percent retained scores of either VR or LM were observed. The authors concluded that the performances of people over 65 on tests of verbal and spatial memory declined even in the absence of medical problems.

Russell (1988) provided norms that included his rescoring procedure (Russell, 1975) and offered age and education corrected scores for individuals age 20 through the age of 80. The normative study was conducted on 188 "normals" referred for neuropsychological testing after displaying symptoms suggestive of organic impairment who later received a negative neurological examination, and 502 brain damaged subjects. The brain damaged group was composed of a number of specific organic pathologies. The mean age and education of the two groups were not found to differ significantly. Russell (1988) noted that the sample contained a low percentage of women and nonwhites but pointed out that neither variable has been found to result in significant differences with regard to memory. In the study, age and education corrections were derived through linear regression analysis based on the assumption that both age and education are related to memory in a linear fashion. This assumption was made solely for the purpose of simplifying the analysis. The resulting age and education corrections to be used for each of the six scores yielded by the Russell (1975) revision of the WMS and a table of corrections for use with the RWMS were provided for subjects age 20-80 and above. One problem with this study is that it does not provide normative data for the original WMS subtest scores. All of the scores were found to predict the diagnostic categories of the subjects with greater than 50% accuracy. The percentages of correct predictions made from a subtest score ranged from 57-76.

In the discussion of the results, Russell (1988) argued against the use of the WMS for differentiating brain damaged individuals from normals. His argument is based on the observation that a number of the "normal" subjects received a RWMS in the slightly impaired range. This argument overlooks the fact that these "normal" subjects were originally referred for neuropsychological testing because their functioning was impaired enough to indicate the possible presence of organic pathology. Therefore, it is unrealistic to expect these subjects to perform well within the normal range of neurological functioning. Additionally, Russell (1988) failed to discuss the fact that the WMS is seldom used in isolation to differentiate brain damaged from normal individuals, and must be evaluated for its usefulness with in the context of a more comprehensive neuropsychological test battery.

In light of the criticism of the WMS, a revised version was made available (Weschler, 1987). The Weschler Memory Scale -Revised (WMS-R) addressed the criticism that the WMS was largely a test of verbal memory by adding three nonverbal subtests (Herman, 1988). The delayed recall procedure outlined by Russell (1975) was adopted and extended to the visual and verbal paired associates test II of the WMS-R as well. The WMS-R yields three separate scores: a) Attention/Concentration b) General Memory and c) Delayed Recall (Weschler, 1987). The orientation and personal and current information subtests of the WMS-R were excluded from the overall scores. Although this revised test increases the nonverbal and delayed recall components of the WMS, it does not bring the measure up to date with current information processing theories of memory. It neglects to include a recognition or cued recall component that would allow a distinction to be made between storage and retrieval deficits. Furthermore, it does not provide any qualitative information about the memory of the examines, such as the number of perseverative or intrusive errors made. It has been suggested that although researchers have consistently called for the development of a new memory scale based on the current memory theories, many clinicians will continue to use the original WMS and base their inferences on the norms available for that test (D'Elia et al., 1989).

A review of the normative studies available for use with both the WMS and Russell modifications and the WMS-R was conducted (D'Elia & Satz, 1989). The paper evaluated the available normative studies according to a set of criteria and found that a number of satisfactory studies are now available for use with the WMS (Abikoff, Alvir, & Hong, 1987; D'Elia et al., 1989; Haaland, et al., 1983; Ivinson, 1986; and Trahan, Qintana, Goethe, & Willingham, 1988). The review did not include Russell's (1988) extensive renorming of his procedure. The paper concludes that there are no satisfactory norms for younger children on measures of delayed recall available and that the norms for delayed visual reproduction and logical memory are limited. The author's criticized both the WMS and WMS-R for failing to provide a recognition component so that problems related to encoding can be differentiated from retrieval problems. The review concludes that the WMS-R provides adequate normative data for a number of groups but that the normative data for individuals 18-19, 25-34, and 45-54 was derived through estimation and is not based on data collected from subjects within these age groups. The authors suggest that the norms for these age groups should not be relied upon. Additionally, no normative studies were reported that included children under 16 on the WMS-R. The original norms (Weschler, 1945) and Cauthen's (1977) norms are recommended for use with the WMS when subjects are between the ages of 20 and 65 (D'Elia et al., 1989).

The early studies conducted on the WMS were fraught with problems and were based on the false assumptions that memory is a unitary function and brain damage is a unitary process. These studies employed poorly defined, heterogeneous groups of neurological and psychiatric patients. Additionally, many of the studies failed to control for IQ and education. The studies also assumed that psychiatric disorders involved no impairment in neurological functioning. When the theories of memory functioning and brain damage began to change, many clinicians adopted the modified procedure for administrating the WMS (Russell, 1975). However, relatively few studies that employed both well defined groups and Russell's (1975) procedure have been conducted. Therefore, the usefulness of this new constellation of scores in yielding patterns of memory performance helpful in differentiating between various neurological and psychiatric groups has not been thoroughly investigated. Additionally, with the exception of Haaland et al. (1983), those studies that did utilize this modified procedure failed to detail the administration and scoring procedures used.

In a comparison of the memory performances of patients with Huntington's Disease and Korsakoff's Syndrome, the groups were found to differ significantly in their recall, recognition, and acquisition of information, although, there was no significant difference in their MQ's (Butters, 1984). This comparison did not include studies that had employed the Russell (1975) modifications. The investigation relied primarily on various individual tests of memory functioning. The study concludes that the MQ and other composite memory scores are of little utility in describing the nature of memory impairments. Although the need for more detailed, qualitative analysis of memory measures is recognized, it is not too late to apply them to the Russell (1975) revision of the WMS.

Altepeter, Adams, Buchanan, & Buck (1987) used the Russell (1975) procedure in comparing the effectiveness of the WMS and The Luria Memory Words Test in differentiating neurologically impaired groups from controls. The study concluded that 72% of the subjects could be correctly identified based on a regression equation that utilized five variables from the WMS; Inf, LM delayed recall, VR delayed recall, AL, and ML. The regression equation resulted in a false positive rate of 17% but this could be reduced to six percent when the WMS variables were combined with those of the Luria Memory Words Test (Altepeter et al., 1990).

Current studies are a vast improvement over the research conducted up until the late 1970's. Technological advances in equipment for the diagnosis of organicity has greatly improved the validity of an organic diagnosis and have made it possible to use highly homogeneous groups. However, several methodological problems still exist. Researchers typically fail to discuss the qualifications of the examiners used to administer the test. More importantly, no discussion of any measures taken to insure inter-examiner reliability is provided. The studies also fail to provide a reliability check for the diagnosis given. The studies that employ Russell's (1975) modified procedure are especially problematic. Some examiners score the logical memory subtest for verbatim recall and others for semantic recall. The instructions to the subject tend to vary with the scoring method used. Although this is important information for replicating a study, the method used is not described in the methodology.

With each development in the research on memory it becomes clear that memory processes are extremely complex. In the 1960's and 1970's it became accepted that memory was far too complex to be described with a single score such as a memory quotient. Attention then began to turn toward individual subtest scores. However, the argument can be made that even subtest scores overshadow important qualitative differences in memory functioning.

In an article by Butters (1984) this problem was addressed for amnesic and demented patients. Butters demonstrated that although the overall performances of the three groups on the WMS Russell revision were not significantly different, qualitative differences that significantly differentiated between the groups could be found. In recent years an increasing amount of attention has been given to the types of memory errors made. The most commonly discussed errors are perseverative, where the subject continues to respond with the same answer either within or between items, and intrusive errors, where something associated to the answer in the subject's mind is offered as part of the response. By considering the types of errors made in response to an item, valuable information about the possible origins of the memory impairment can be obtained. For example; if a subject responds to the question "Who is the president of the United States?" with the name of a former president, this may indicate that past memories have been left intact, while further evaluation of the acquisition phase of memory is needed. If a subject responded to the same question by saying that he knows it starts with a B and that he used to be a vice president, this would indicate that the subject has acquired the information but is having difficulty retrieving it, and further evaluation of recognition memory may be needed. Therefore, when only the subtest scores are considered, a great deal of valuable information is lost. Undoubtedly, experienced neuropsychologists use this qualitative information to assist in diagnosis and clinical recommendations. However, such detailed discussions in the memory literature are rare. This study will examine the differences in the types of errors made and present any prevalent patterns found within a diagnostic group. As a detailed examination of every item for every subject in this study would be cumbersome and time consuming, the qualitative analysis will be limited to those subtests found to significantly differentiate between the diagnostic groups. It is hoped that this will increase the understanding of the memory deficits involved and help to generate hypotheses for future research.

The present study was designed to examine the utility of using the WMS administered according to Russell's (1975) modified procedure to differentiate between well defined groups of neurologically and psuedoneurologically impaired subjects: Closed head injury (CHI), Cerebrovascular disease - left hemisphere (CV-LH), Cerebrovascular disease - right hemisphere (CV-RH), Dementia (DEM), and Pseudoneurological (PSDEM). The study also attempts to address the methodological problems of past studies by using a standard format in administering and scoring the Russell (1975) modifications and performing reliability checks on the diagnosis of those subjects included in the study. A function was derived based on the variables which best differentiated between the diagnostic groups. This function was then used to predict group membership for a second set of subjects. The study carefully screened each subject and employed reliability checks on the diagnosis of each subject to increase the homogeneity of the groups. Age and education were partially controlled for through approximate matching and then treated as experimental variables. Although the WMS-R provides additional information, the testing laboratory where this study was conducted, and many others, would prefer to continue using the WMS with Russell's modifications. This reluctance is due in part to the greater experience the staff has in administering and interpreting the results of the WMS Russell revision. It is also influenced by the archival data that exists on the WMS and the problems that were not corrected by the WMS-R. The WMS-R does not provide adequate norms for all age groups. Additionally, because the WMS-R failed to include recognition and serial recall components, supplementary tests are still needed to assess memory more thoroughly. In light of the belief that many clinicians will continue to use the WMS Russell revision, further study is needed to assess its utility as a diagnostic tool. It is hypothesized that well defined and qualitatively different neurologically and pseudoneurologically impaired groups will exhibit distinctively different patterns of performance on the WMS and that a function that will predict group membership on this basis can be derived. Although significant between-group differences in WMS performance are expected, no specific predictions of group performance were made.

CHAPTER III

EXPERIMENT 1

Method

Subjects

The subjects used in this study were individuals selected from the files of the Psychological Assessment Laboratory of the Oklahoma Health Sciences Center. The subjects were selected on the basis of a biomedical diagnosis of diffuse closed head injury (CHI), cerebrovascular disease strongly lateralized to the left hemisphere (CV-LH), cerebrovascular disease strongly lateralized to the right hemisphere (CV-RH), dementia (Alzheimer's or Picks type; DEM), or pseudodementia (subjects that manifested neurological symptoms which were not supported by the biomedical test results; PSDEM). The biomedical diagnosis involved positive results on one or more of the following; Computerized Axial Tomography (CAT Scan), multiple electroencephalograms (EEG), Magnetic Resonance Imaging (MRI), angiogram, and various other tests of diagnostic imaging. Only those diagnoses that could be made with strong confidence were included in this study. Subjects that fit the criteria for more than one diagnostic group were not included. Additionally, subjects with a history of seizure disorder or drug and alcohol abuse were excluded from the study because of the neurological effects of seizures and drug abuse. Each subject's biomedical information was blindly reviewed by the researcher and only those subjects for whom an identical diagnosis was made were included in the study. Subjects that met the above criteria but did not complete the Wechsler Memory Scale (WMS) according to the Russell revision were excluded from the study. Additionally, only subjects between the ages of 20 and 65 were selected in an effort to reduce age related performance differences.

Fifty one subjects met the criteria for inclusion in the study. There were 8 males and 3 females in the CHI group, 3 males and 5 females in the CV-LH, 4 females and 4 males in the CV-RH group, 6 males and 6 females in the DEM group, and 8 males and 4 females in the PSDEM group. Because of a statistical need to equalize group size for the proposed discriminant function procedure, 8 subjects were selected from the CHI, DEM, and PSDEM groups on the basis of approximate matching for age, education, and sex. The DEM group was used as the referent group for matching as it was the most restricted with regard to age and education. The CHI group was also somewhat restricted and was composed primarily of subjects in their 20's and 30's. The 8 youngest DEM subjects were selected from the initial 12. Each DEM subject was then compared to those in the remaining 4 groups and the subject in each group closest in age was chosen. When several subjects were equally matched for age,

education and sex were then matched for in making the selection. Subjects were first approximately matched for age then education and finally sex. The matching was approximate because the clinical specificity of the groups made it impossible to match the subjects precisely for these variables. The inclusion of both Dem and CHI created the biggest problem in controlling for age.
Additionally, only 8 subjects met the criteria for the CV-LH and CV-RH groups making this procedure meaningless for these two groups. The remaining 11 subjects were included in the second experiment. The characteristics of the subject groups are summarized in Table 1.

Insert Table 1 about here

In a study by Albikoff et al. (1987) the impact of age and education on WMS scores was assessed, and both factors were found to be related to WMS scores in a nonlinear fashion. Age and education were also judged to be confounded variables with education accounting for a large percentage of the variance in WMS scores. However, even with education effects removed, age was significantly correlated to recall though the coefficients were much lower. In this study, the effects of age, education, and sex were controlled for to the degree possible by the approximate matching procedure discussed above.

Procedure

Table 1

Descriptive Statistics Experiment 1

		AGE		EDUC		M	F
Group	n	Mean	SD	Mean	SD		
CHI	8	33.37	10.79	12.31	1.75	5	3
CV-LH	8	46.12	12.15	13.50	1.87	3	5
CV-RH	8	44.62	10.84	13.50	4.527	4	4
DEM	8	53.75	4.35	13.62	1.58	4	4
PSDEM	8	41.50	14.06	13.68	2.97	5	3

Note: See Table 2 for abbreviations.

Each of the subjects in this study were referred to the Psychological Assessment Laboratory at the Oklahoma Health Sciences Center for difficulties suggestive of neurological impairment. In this procedure, subjects were all given the WMS according to the modified procedure outlined by Russell (1975).

The subject is instructed to repeat everything that is read to him. A half-credit scoring method that allows for semantic and verbatim recall was utilized. For a more detailed discussion of the administration and scoring procedures used see Appendix A. The WMS was administered within the context of a standard neuropsychological assessment that included the Wechsler Adult Intelligence Scale - Revised, the complete Halstead-Reitan Neuropsychological Test Battery for Adults, Trail Making Test, and Luria Memory Words. The test batteries were administered and supervised by experienced psychometrists and neuropsychologists employed by the Oklahoma Health Sciences Center. During the period between the first testing for figural and semantic memory and the 30 minute delayed recall of these two subtests, the Trails Making Test or a similar test that did not involve a memory component was administered. Additionally, 7 variables of interest from the WMS were also coded. They include: Counting by threes (3's), AL easy, Al Hard, and the scores on the initial drawing of each individual figure (A, B, C1, & C2). These variables were selected based on an article by Howard (1950). Age and Educ were also treated as experimental variables and included in the GLM procedure. A list of abbreviations is provided in Table 2.

Insert Table 2 about here

Statistics

Univariate F tests were first conducted on the entire subject population. Tukey's (HSD) Test of pairwise comparisons was also conducted. These statistics were then computed on each WMS subtest raw score: Inf, O, MC, LM, MS, VR, & AL. They were also computed on each additional score yielded by the Russell procedure: LM delayed, VR delayed, percent retained LM (% LM), percent retained VR (% VR), and Russell's WMS average impairment rating (RWMS), and the additional variables described in the procedure: 3's, AL easy, AL hard, and the scores on the initial drawings of each figure: A, B, C1, and C2 in addition to age. A subsequent Stepwise Discriminate Function analysis was then conducted. For this procedure equal group size is optimal. Group size for this procedure was therefore determined by the number of subjects in the smallest group. As the number of variables entered into a discriminate function analysis should not exceed the number of subjects per group, the results of the Univariate F tests were used to determine which variables were entered into this analysis. Only those variables with F tests significant at the .05 level were included in the analysis. A Canonical Discriminate function

Table 2

List of Abbreviations

Abbreviation	Variable or Group Name
CHI	Closed head injury
CV-LH	Cerebral vascular accident - left hemisphere
CV-RH	Cerebral vascular accident - right hemisphere
DEM	Dementia
PSDEM	Neurologic symptoms - no biomedical evidence
INF	Information subtest
OR	Orientation subtest
MC	Mental Control subtest
LM	Logical Memory subtest
DIG	Memory for Digits subtest
VR	Visual Reproduction subtest
AL	Associate Learning subtest
LM delayed	Logical Memory, delayed recall
VR	Visual Reproduction, delayed recall
VR%	Visual Reproduction, percent retained
RWMS	Russell's WMS average impirment rating
WMQ	Wechsler Memory Quotient
FIGA	Recall on Figure A
FIGB	Recall on Figure B
FIGC1	Recall on Figure C1
FIGC2	Recall on Figure C2
AL easy	Associate Learning, easy pairs
AL hard	Associate Learning, hard pairs
3's	Mental Control, counting by 3's

analysis was then conducted to evaluate the validity of the resulting function. A subsequent qualitative analysis was planned for those variables which entered into the discriminate function if the additional variables included in the analysis were found significantly more useful differentiating between the groups than the overall subtest scores. The additional variables did not prove to provide better information; therefore, a further qualitative analysis was not conducted.

Results

A General Linear Models Procedure (GLM) was used to conduct Univariate F tests on each of the WMS Russell Revision variables included in the study. The Univariate F tests for each of the variables was significant at or above the .05 level with the exception of EDUC, INF, LM, LM delayed, and %LM. A complete listing of the univariate test statistics is provided in Table 3.

Insert Table 3 about here

In the second phase of the analysis, variables were selected for entrance into a Stepwise Discriminant Function Analysis. The number of variables entered into the equation was restricted to seven (the number of subjects with complete data present in the smallest group) for the statistical reasons previously discussed. The number of variables with Univariate F tests significant at the .05

Table 3

Summary of the Univariate Tests Below the .05 Level of Significance

Variable	n	F	p
AGE	40	3.31	.021
WMQ	40	4.67	.004
VR	40	9.49	.0001
FIGA	40	2.91	.0355
FIGB	40	4.37	.0057
FIGC1	40	13.70	.0001
FIGC2	40	6.46	.0005
VR delayed	40	13.40	.0001
%VR	40	10.07	.0001
RWMS	38	11.49	.0001
DIG	36	4.09	.0089
OR	39	5.46	.0017
AL	39	4.48	.0051
3's	40	3.14	.0262
AL easy	40	2.87	.0369
AL hard	40	3.64	.0140
MC	39	2.87	.0376

Note: See Table 2 for abreviations

level of significance exceeded the number of variabls appropriate for entrance into the Stepwise Discriminant Function Analysis. Therefore, the variables were selected based on the p value of the Univariate F tests and the number of subjects for which the score was available. The discriminant function procedures automatically exclude all data for any subject whose record is not complete for the variables included in the statistical analysis. Therefore, although RWMS, and MS had significant F tests, they were not included in the Stepwise Discriminant Function analysis due to the number of subjects for which these scores were not available. Several other variables also had significant F tests but were excluded because of their redundancy. Age was included in this analysis to absorb any of the variance among the remaining scores that may be due to age effects.

The set of variables entered into the Stepwise Discriminant Function analysis were thus VR delayed, VR, O, AL, MQ, MC, and AGE. The resulting Stepwise Discriminant Function included VR delayed, VR, O, and MC all of which had F statistics that reached the .05 level of significance for entrance into the equation (Wilk's L=.15, F (16, 92.28)=5.045, p=.0001). The VR delayed score alone recieved an F statistic of 12.72 and p=.0001, thus accounting for a large portion of the variance between the groups. The F statistics for each of these variables are included in Table 4. A subsequent Canonical Discriminant Function analysis was then performed and resulted in a Wilk's L=.12, F (24, 98.89)=3.54, and p=.0001. The Cannonical Discriminant

Function takes each subject in Experiment 1 individually and applies the derived function to test its' ability to predict group membership. This analysis serves as a measure of integrety for the resulting Stepwise Discriminant Function. However, this procedure calculates constants for each of the variables originally entered into the Stepwise Discriminant Function procedure whether or not they were included in the resulting stepwise discriminate function.

Insert Table 4 about here

Tukey's (HSD) Test of pairwise comparisons was also conducted for all variables with significant Univariate F Tests. The results for those variables chosen for entrance into the Stepwise Discriminant Function analysis are reported in Table 5. Additionally, the descriptive statistics for subscale scores with significant Univariate F tests are also reported in Table 6.

Insert Table 5 & 6 about here

EXPERIMENT 2

Method

Subjects

The subjects consisted of 16 individuals referred to the neuropsychological assessment laboratory for testing.

Table 4

Stepwise Discriminant Function Analysis

Step	Variable	F Statistic	Prob > F
1	VR delayed	12.715	.0001
2	VR	3.338	.0215
3	OR	2.929	.0365
4	MC	2.807	.0432

Note: See Table 2 for abreviations

Table 5

Tukey's Studentized Range (HSD) Test

Pairwise Comparisons	AGE	MQ	VR	VR2	OR	MC	AL	%VR	RWMS	MS
	Variables									
CHI/CV-LH				***						
CHI/CV-RH								***		
CHI/DEM	***						***			
CHI/PSDEM		***		***				***	***	
CV-LH/CV-RH			***							
CV-LH/DEM			***		***			***	***	
CV-LH/PSDEM						***				***
CV-RH/DEM				***	***			***	***	
CV-RH/PSDEM			***	***						
DEM/PSDEM		***	***	***			***	***	***	***

Note: *** indicates that the pairwise comparison was significant at the .05 level of significance with respect to the corresponding variable. Those group pairs for which no *** appears under a given variable, had means that were not significantly different with respect to that variable.
See Table 2 for abbreviations.

Table 6

Descriptive Statistics for each Variable with a Significant Univariate F Tests in Experiment 1

	CHI			CV-LH			CV-RH			DEM			PSDEM		
Var	n	m	SD	n	m	SD	n	m	SD	n	m	SD	n	m	SD
WMQ	8	86.3	12.7	8	99.4	25.8	8	93.4	15.4	8	78.8	11.7	8	116	21.4
VR	8	6.3	2.7	8	9.1	3.4	8	5.3	2.4	8	2.4	1.9	8	10.1	3.5
VR-2	8	2.3	2.4	8	7.1	3.3	7	5.0	2.16	8	0.8	1.4	8	9.6	3.8
%VR	8	39.7	40.8	8	78.7	19.8	7	89.6	18.7	8	20.0	38.5	8	92.3	11.1
RWMS	8	3.4	0.8	8	2.2	0.9	6	2.5	0.8	8	3.9	0.9	8	1.3	0.9
DIG	7	10.0	1.8	7	7.9	2.7	7	10.7	1.6	8	8.5	2.3	7	11.9	2.2
OR	7	3.8	0.7	8	4.8	0.5	8	4.8	0.5	8	3.4	1.2	8	4.4	0.5
AL	7	15.5	4.7	8	12.5	5.0	8	12.3	4.2	8	8.3	3.5	8	16.5	3.8
MC	7	5.6	1.5	8	4.4	2.7	8	5.6	2.3	8	4.8	1.5	8	7.4	1.2

Note: See Table 2 for abreviations

Eleven of the subjects were carried over from the original subject pool. The remaining subjects were collected according to the same criteria outlined in Experiment 1 with the exception of age. The age criteria was extended to include two additional subjects in the CV-LH, and DEM groups, age 66 and 67 respectively. There were initially 4 subjects in the CHI group, 2 in the CV-LH group, 1 in the CV-RH group, 5 in the DEM group, and 4 in the PSDEM group. The descriptive statistics are presented in Table 7. As Discriminant Function Analysis rejects all subjects with any missing data, only 11 subjects were included in this analysis. One subject from each group was excluded from the analysis.

Insert Table 7 about here

Procedure

The same procedure for data collection described in Experiment 1 was applied. However, subjects were not separated into diagnostic groups. They were grouped by the discriminant function yielded in Experiment 1.

Statistics

This analysis consisted of the application of the Discriminant Function resulting from Experiment 1 to the new group of subjects. The Stepwise Discriminant Function procedure does not allow for the entry of a test data set.

Table 7

Descriptive Statistics Experiment 2

		AGE		EDUC	
Group	n	Mean	SD	Mean	SD
CHI	4	24.00	3.46	13.33	2.08
CV-LH	1	53.00	...	15.00	...
CV-RH	0	...	...	...	...
DEM	4	63.25	2.36	16.00	3.26
PSDEM	4	34.75	13.89	11.50	1.73

Note: See Table 2 for abbreviations.

In order to correct for this problem, the four variables selected by Stepwise Discriminant Function Analysis (VR delayed, VR, MC, & O) were put into a Discriminant Function Analysis. This procedure allows for the entrance of a test data set. This discriminant function was then applied to the new set of subjects. Therefore, it serves only as an approximate test of the validity of the Stepwise Discriminant Function obtained in Experiment 1.

Results

The discriminant function performed at a below chance level in correctly categorizing the subjects into their respective groups. As discriminant function analysis rejects all observations for any subjects with missing data, one subject from the CHI group, one from the CV-LH group, one from the CV-RH group, one from the DEM group, and one subject from the PSDEM group were excluded from the analysis. This resulted in no classification data being available for the CV-RH group. The percentage of subjects correctly classified were; CHI group 33.33% (1 of 3), CV-LH group 0% (0 of 1), CV-RH group not available, DEM group 50% (2 of 4), and the PSDEM group 0% (0 of 3). The Univariate *F* tests for each group were also calculated. The results are provided in Table 8. Three of the four variables included in the stepwise discriminant function failed to achieve significance at the .05 level of significance. However, VR delayed was significant at the .02 level.

Insert Table 8 about here

Table 8

Summary of the Univariate Tests Above the .05 Level of Significance for Experiment 2

Variable	n	F	p
AGE	11	13.32	.0018
WMQ	11	0.10	.96
VR	11	1.49	.29
FIGA	11	2.21	.13
FIGB	11	2.83	.078
FIGC1	11	1.22	.36
FIGC2	11	1.12	.39
VR delayed	11	5.79	.02
%VR	11	10.72	.0036
RWMS	11	4.71	.04
DIG	11	.70	.58
OR	11	.90	.49
AL	11	3.89	.06
3's	11	1.42	.29
AL easy	11	1.63	.24
AL hard	11	6.00	.008
MC	11	3.66	.071

Note: See Table 2 for abreviations

CHAPTER IV

DISCUSSION

Univariate F Tests Experiment 1

The results of Experiment 1 suggests that the WMS Russell Revision is useful in differentiating between groups of neurologically and pseudoneurologically impaired subjects. The Univariate F tests found the groups to differ in their performance on nine out of the 13 scores obtained from the Russell administration of the WMS at the .05 level of significance. The additional WMS values included in the analysis were also significant at the .05 level of significance. These findings are not consistent with earlier studies (Cohen, 1950; Parker, 1957; & Shontz, 1957) which found the WMS ineffective at differentiating among various diagnostic groups. The results of this study suggests that past failures to find the WMS to be useful in differentiating among various diagnostic group were largely a function of the flawed methodologies and broad diagnostic categories characteristic of these earlier studies. The results support Prigatano's (1978) conclusion that the WMS is a useful instrument and suggest that although the WMS is not fully consistent with current memory theory, it is effective in the assessment of memory functioning in impaired populations.

The usefulness of the Russell modified scoring procedure is illustrated by the finding that many of these scores received the F statistics of the highest magnitude. This suggests that the Russell Revision provides additional information useful in distinguishing among the five diagnostic groups included in the study.

The Tukey (HSD) pairwise comparisons reveal that VR is the most useful in distinguishing the PSDEM and DEM subjects from the other diagnostic groups. The Tukey comparisons also indicate that OR is most useful in differentiating the DEM subjects from the CV-LH and CV-RH group while MC was useful in distinguishing the PSDEM subjects from the CV-LH subjects. The failure of past studies to find these variables as significant may be explained by the fact that past studies have not employed groups of subjects with cerebral vascular accidents.

Discriminant Function Analysis

The results of the Stepwise Discriminant Function analysis and Canonical Discriminant Function analysis support the hypothesis that a function useful in predicting group membership can be derived from the WMS Russell Revision subtest scores. These findings are consistent with the those of Howard (1954), Bachrach & Mintz (1974), and Altepeter et al. (1987) which found the WMS is useful in differentiating among various diagnostic groups when factors such as age, education, and/or severity of injury are controlled for. The Canonical Discriminant Function analysis found that group membership could be predicted by the resulting Stepwise Discriminant Function. The PSDEM group was differentiated from the remaining groups on the basis of age, MQ, VR, VR delayed, AL, MC, and OR with 87.5% accuracy when education was controlled for. This finding is of particular interest as distinguishing PSDEM subjects from neurologically impaired subjects is the most important and difficult distinction neuropsychologist are asked to make.

The resulting Stepwise Discriminant Function included VR, VR delayed, MC, and OR. The VR delayed score alone resulted in the correct classification of subjects at the .0001 level of significance and accounted for a large portion of the variance among the groups. The second most significant variable was VR. This finding is consistent with the results reported by Bachrach and Mintz (1974) that found the VR score to result in 89% correct classification of subjects in to their respective groups. The PSDEM group performed better on this subtest than the remaining groups while the DEM groups were the most impaired.

The significance of the VR scores is likely due to the fact that the ability to copy figures from memory is seldom impaired in subjects with no neurologic impairment. The degree of impairment on VR measures would be expected to increase with the severity and diffuseness of the neurologic involvement. However, when damage is lateralized to the right hemisphere which is largely responsible for visual spatial abilities, more severe impairment of the VR scores is expected. The group means are consistent with what would be expected. The CV-LH subjects with strongly lateralized damage were less impaired on measures of VR than the CHI and DEM groups. The CV-RH group was less impaired than the DEM group, but more impaired than the CHI group. The inclusion of the variables OR and MC are not consistent with past studies that have not reported these variables as useful in distinguishing between various diagnostic groups.

Applied Discriminant Function Analysis

The results of Experiment 2 ,however, do not support the usefulness of the Stepwise Discriminant Function derived in Experiment 1. The function was found to perform at the below chance level in predicting group membership for a new group of subjects. However, there were several limitations placed on this phase of the analysis which must be considered in interpreting this result. Due to the small number of subjects available which met the criterion for inclusion in a diagnostic group, it was not possible to approximately match the subjects for age or education. A comparison of the descriptive statistics for both subject pools reveals that the subjects included in the second experiment were on the extreme ends of the age ranges for their respective groups. Additionally, as the Stepwise Discriminant Function procedure does not allow for the inclusion of a test data set, a procedure that would approximate such results was utilized. This involved inserting the four variables included in the Stepwise Discriminant Function analysis into a Discriminant Function analysis and applying these results to a second subject pool. One limitation of this procedure was that it did not allow for the inclusion of age. In the first experiment, age was entered in to the discriminant function procedure to absorb its share of the variance. The Univariate F tests performed on the second set of subjects found groups to differ with respect to age at the .0018 level of significance. Therefore, age has likely confounded the results.

General Discussion

The results of this study suggest that the WMS Russell Revision is useful in differentiating between well defined groups of neurologically impaired subjects when the effects of age and education are controlled for. The study also finds this instrument to be very useful in distinguishing pseudoneurologic individuals from subjects with various neurologic impairments. The hypothesis that individuals with neurological symptoms of various eitiologies would yield profiles characteristic of their neurological conditions was confirmed. However, it was not possible to derive a useful formula from this data for the prediction of group membership on the basis of WMS scores with discriminant function procedures. The significance of the Univariate F tests suggests that the failure to derive a reliable formula was likely a factor of the statistical limitations inherent in multivariate statistics. A profile analyisis approach to MANOVA or a pattern analysis based on the Univariate F tests is suggested for future research. These results support the continued use of the WMS Russell Revision and the need for further research on the complete profile of scores. The importance of controlling for age and education further illustrate the need for the collection of norms on all scores resulting from this measure and the derivation of age and education and education corrections for each score.

The inclusion of additional variables derived from the overall subtest scores provided useful information on the specificity of the WMS scores. Although the Univariate F tests conducted on these variables were significant, they were no more significant than the overall subtest scores. This finding is consistent with similar analyses conducted in earlier studies (Howard, 1950 & 1954). These results suggest that the examination of individual items on the WMS Russell Revision is not necessary given the usefulness of the subtest scores themselves.

The Univariate F tests conducted on the subjects in Experiment 2 suggest that VR delayed is reliably useful in differentiating among CHI, CV-LH, CV-RH, DEM, and PSDEM groups. This replication adds further support to the results of Experiment 1 which found VR delayed to have a large F statistic and account for a large portion of the variance among groups. The replication of this finding suggests that VR delayed reliably differentiates between these five diagnostic groups. The other three variables present in the stepwise discriminant function (VR, MC, & O) did not produce significant Univariate F tests for the second group of subjects. However, MC was within the trend level of significance ($p < .10$). It is important to note that several other variables did achieve significant Univariate F tests in both experiments. These variables were age, VR%, RWMS, AL, & AL hard. With the exception of VR%, these variables were not included because there were patients with missing data on these measures. Visual Reproduction % was not included because each of the VR variables was significant and VR% has been described as the least meaningful. The comparisons of the two sets of Univariate F tests suggests that other variables may have produced a more reliable stepwise discriminant function. These results again illustrate the limitations of the discriminant function procedures. The need to limit the number of variables entered in to the equation to the number of subjects for each group can require the researcher to make nonstatistical decisions as to which variables to include in the analysis. Ideally the statistical procedure would make these decisions. However, the number of subjects required for each group and the clinical specificity of the groups makes this impractical. Additionally, the fact that discriminant function procedures omit any subject from the analysis with missing data on any one variable further complicates this research. The Univariate F test results certainly make a strong argument for the future use of a Pattern Analysis for the interpretation and application of WMS Russell revision scores in differential diagnosis. Experiment 2 suggest that VR delayed is reliably useful in differentiating between CHI, CV-LH, CV-RH, DEM, and PSDEM groups.

The present study has attempted to correct many of the methodological problems present in previous studies. However, the study was limited by the small number of subjects available and the resulting need to limit the variables entered in to the Stepwise Discriminant Function analysis. This study clearly demonstrated the usefulness of the WMS Russell Revision in differentiating between groups, although a discriminant function approach did not prove useful. Several scores emerged as reliably differentiating. These variables were; VR delayed, %VR, RWMS, and AL. The reliability of these scores warrent further research on how these variables can be applied in differential diagnosis. Future research should also focus on developing adequate norms and age and education corrections for the WMS Russell Revision subtest scores. Once the age and education corrections become available, further studies conducted to assess the ability of the corrected scores to differentiate between well defined diagnostic groups will be needed.

REFERENCES

Abikoff, H., Alvir, J., & Hong, G. (1987). Logical Memory subtest of the Wechsler Memory Scale: Age and education norms and alternative form reliability of two scoring systems. Journal of Clinical and Experimental Neuropsychology, 9, 435-448.

Altepeter, T. S., Adams, R. L., Buchanan, W. L., & Buck, P. (1990). Luria Memory Words Test and Wechsler Memory Scale: Comparison of utility in discriminating neurologically impaired from controls. Journal of Clinical Psychology, 46, 190-193.

Bachrach, H., & Mintz, J. (1974). The Wechsler Memory Scale as a tool for the detection of mild cerebral dysfunction. Journal of Clinical Psychology, 30, 58-60.

Bartlett, F. C. (1932). Remembering. Cambridge: Cambridge University Press.

Black, F. W. (1973). Cognitive and memory performance in subjects with brain damage secondary to penetrating missile wounds and closed head injury. Journal of Clinical, Psychology, 29, 441-442.

Butters, N. (1984). The clinical aspects of memory disorders: Contributions from experimental studies of amnesia and dementia. Journal of Clinical Neuropsychology, 6, 17-36.

Cauthen, N. R. (1977). Extension of the Wechsler Memory Scale norms to older age groups. Journal of Clinical Psychology, 33, 208-212.

Cohen, J. (1950). Wechsler Memory Scale performance of psychoneurotic, organic, and schizophrenic groups. Journal of Consulting Psychology, 14, 371-375.

D'Ellia, L. , Satz, P., & Schretlen, D. (1989). Wechsler Memory Scale: A critical appraisal of the normative studies. Journal of Clinical and Experimental Neuropsycholgy, 11, 551-568.

Dujouvne, B. E., Levy, B. I. (1971). The psychometric structure of the Wechsler Memory Scale. Journal of Clinical Psychology, 27, 351-354.

Ebbinghaus, H. (1913). Memory. New York: Teachers College Press.

Erickson, R. C., & Scott, M. L. (1977). Clinical memory testing: A review. Psychological Bulletin, 84, 1130-1149.

Guilford, J. P. (1967). The nature of human intelligence. New York: McGraw-Hill.

Haaland, K. Y., Linn, R. T., Hunt, W. C., & Goodwin, J. S. (1983). A normative study of Russell's variant of the Wechsler Memory Scale in a healthy elderly population. Journal of Genetic Psychology, 109, 135-145.

Heaton, R. K., Baade, L. E., Johnson, K. (1978). Neuropsychological Test Results Associated With Psychiatric Disorders in Adults. Psychological Bulletin, 85, 141-162.

Herman, D. O., & Young, L. L. (1987, August). Development of the Wechsler Memory Scale - Revised. The Clinical

Neuropsychologist, 2, 102-106.

Howard, A. R. (1950). Diagnostic value of the Wechsler Memory Scale with selected groups of institutionalized patients. Journal of Consulting Psychology, 14, 376-380.

Howard, A. R. (1954). Further validation studies of the Wechsler Memory Scale. Journal of Clinical Psychology, 10, 164-167.

Hulica, J. M. (1966). Age differences in Wechsler Memory Scale scores. Journal Genetic Psychology, 109, 135-145.

Ivinskis, A., Allren, S., & Shaw, E. (1971). An extention of Wechsler Memory Scale norms to lower age groups. Journal of Clinical Psychology, 27, 354-357.

Kear-Colwell, J. J. (1973). The structure of the Wechsler Memory Scale. Journal of Clinical Psychology, 27, 351-354.

Kesner, R. (1973). A neural system analysis of memory storage and retrieval. Psychological Bulletin, 80, 177-203.

Kijajic, I. (1975). Wechsler Memory Scale indices of brain pathology. Journal of Clinical Psychology, 31, 698-701.

Kimura, D. (1963). Right temporal lobe damage. Archives of Neurology, 8, 264-271.

Larabee, G. J., Kane, R. L., Schuck, J. R., & Francis, D. J. (1985). Construct validity of various memory testing proceedures. Journal of Clinical and Experimental Neuropsychology, 7, 239-250.

Milner, B. (1968). Visual recognition and recall after temporal lobe excision in man. Neuropsychologia, 6, 191-209.

Milner, B. (1970). Memory and the medial temporal regions of the brain. In K. H. Pribam & D. E. Broadbent (Eds.), Biology of memory. New York: Academic Press.

Parker, J. W. (1957). The validity of some current tests for organicity. Journal of Consulting Psychology, 21, 425-428.

Prigatano, G. P. (1978). Wechsler Memory Scale: A selective review of the literature. Journal of Clinical Psychology, 34, 816-832.

Rappaport, D. (1945). Diagnostic psychological testing. Chicago: Year Book Publishers.

Russell, E. W. (1975). A multiple scoring method for the assesment of complex memory functions. Journal of Consulting and Clinical Psychology, 43, 800-809.

Russell, E. W. (1988). Renorming Russell's version of the Wechsler Memory Scale. Journal of Clinical and Experimental Neuropsychology, 10, 235-249.

Shontz, F. C. (1957). Evaluation of intellectual potential in hemiplegic individuals. Journal of Psychology, 13, 267-269.

Wechsler, D. (1944). The measurement of adult intelligence. Baltimore: Wiliams & Wilkins.

Wechsler, D. (1945). A standardized memory scale for clinical use. Journal of Psychology, 19, 87-95.

Wechsler, D. (1987). Wechsler Memory Scale - Revised. New York: The Psychological Corporation.

APPENDIX

RUSSELL METHOD OF THE WECHSLER MEMORY SCALE (USED IN OKLAHOMA)

1. Logical Memory instructions:

First Recall - " I am going to have you repeat a story for me. We will do it once now and, I'll also ask you to recall it "later". I am going to read to you a little section of about 4 or 5 lines, Listen carefully because when I am through I want you to tell me everything I read to you. Are you ready?" (In Oklahoma the order of the stories has been reversed for research purposes). After reading the first story, say, "Now what did I read to you? Tell me everything, and begin at the beginning." (Record verbatim - and score according to the number of ideas - 1 point or 1/2 point - as marked off in the selection.

Delayed Recall - After the first recall a 30 minute delay begins. During the 30 minute delay period the interposed tests should be quite different from either of the memory tests so that contamination and interference do not occur. After the end of the 30 minute dely, the subject is asked to retell the stories as he remembers them. The subject can recall the stories in any order. If the subject cannot remember one or both of the stories at all he is prompted, "Do you remember a story about a washer woman? or Do you remember a story about a ship?" If he then recalls the story, it is scored according to the above instructions except the prompt is not included in the score. At the end of each story it is permissible to say is there anything else that you might remember?

2. Visual Reproduction instructions:

First Recall - At the begining of the visual reproduction say: "I am going to have you reproduce some drawings. I will have you do them once now, and then again "later". I am going to show you the first drawing. You will have 10 seconds to look at it; then I shall take it away and let you draw it from memory as carefully as you can. Don't begin to draw until I say 'go'. Ready?" Expose for 10 seconds. (The 3 design cards are given in order of B, A, C. In Oklahoma we are reversing the order of the first 2 figures for research purposes)

Delayed Recall - The subject is asked to reproduce the designs again after a 30 minute delay. If he does not remember the first design (Fig B), he is given the clue: "Do you remeber a design that looks like squares?" If he does not remember the second design (Fig A), he is given the clue "Do you remember a design that looks like flags?" On the last card that had two designs on it the examiner may say: "Do you remember the card that had 2 designs on it side by side?" If a prompt is given for Figures B or A deduct one point from that figures score. If a prompt is given for Figures C1 and C2, do not deduct a point.

VITA

Mary Ann Scott

Candidate for the Degree of

Doctor of Philosophy

Thesis: WECHSLER MEMORY SCALE RUSSELL REVISION: DIFFERENTIATION OF CLOSED HEAD INJURY, STROKE, DEMENTIA, AND PSEUDODEMENTIA

Major Field: Clinical Psychology

Personal Data: Born in Osceola, Missouri, July 16, 1963,daughter of James and Velma Scott. Married to Mehdi Sarvestani January 1, 1989.

Education: Graduated from Forsyth High School, Forsyth, Missouri in May, 1981. Received a Bachelor of Arts degree from the University of Missouri - Columbia in May, 1986; Received a Master of Science degree from Oklahoma State University, Stillwater, Oklahoma in December, 1988; completed requirements for the Doctor of Philosophy degree from Oklahoma State University, Stillwater, Oklahoma in July, 1992.

Professional Experience: Psychological Assistant, Oklahoma State University Psychological Services Center, September, 1988 to July, 1990; Marriage & Family Clinic June, 1989 to December 1989; Edwin Fair Mental Health Center July, 1989 to June, 1990; Payne County Guidance Center September, 1989 to December 1990 and September 1990 to June 1991. Instructor for General Psychology Correspondence course, Oklahoma State University, September, 1987 to August, 1988. Research Assistant Oklahoma Health Sciences Center Psychological Assessment Lab, June to August, 1988. Teaching Assistant for Abnormal Psychology August, 1988 to May, 1989. Psychology Intern, University of Texas Health Sciences Center - Houston, July, 1991 to June, 1992.