The Development of a Scoring System for an Alternative Form of the Visual Reproduction subtest of the Wechsler Memory Scale - Revised

Daniela Petrov
BSc. (Hons)

A dissertation submitted in partial fulfilment of the requirements for the degree of Doctor of Psychology (Clinical Neuropsychology)

Department of Psychology
Victoria University
Melbourne, Australia

December 2007

DECLARATION

I, Daniela Petrov, declare that the Doctor of Psychology (Clinical Neuropsychology) thesis entitled “Development of a scoring system for an Alternative Form of the Visual Reproduction subtest of the Wechsler Memory Scale – Revised Visual Reproduction subtest” is no more than 40,000 words in length, exclusive of tables, figures, appendices, references and footnotes. This thesis contains no material that has been submitted previously, in whole or in part, for the award of any other academic degree or diploma. Except where otherwise indicated, this thesis is my own work.

Signature: ____________________ Date: ___________

Submitted with emendations: October 2008

ACKNOWLEDGEMENTS

I would like to thank my initial supervisor Dr. Peter Dowling for his wisdom, guidance and support and for generously giving his time to assist me in completely this project.

I would also like to thank my second supervisor Dr. Gerard Kennedy for his timely feedback and editing of drafts.

To my family and friends who have provided me with unconditional love and support throughout this chapter of my life, I thank-you sincerely.

My gratitude extends to the individuals who generously volunteered their time to participate in this study without whom, this project would not have been possible.

TABLE OF CONTENTS

ABSTRACT

The Visual Reproduction subtest of the Wechsler Memory Scale – Revised has been widely used in neuropsychological assessment as a measure of visual memory since its publication in 1987 by Wechsler. Mindful of practice effects in serial assessment, a potential limiting factor in assessing recovery or decline with this subtest is the absence of an alternative form. This study aimed to develop an alternative form and a scoring system for this form. The scoring system was modelled on a revised scoring system for the original version developed by Clark (2000). Additional procedures, namely a cued recall and recognition format, were included in the administration to improve the diagnostic utility of the subtest. Evaluation of the psychometric properties of the Alternative Form, based on the administration of both versions to an unselected adult non-clinical population (n = 44, aged 25-51) revealed high internal reliability, good convergent validity and satisfactory discriminant validity with other verbal and visual memory tests. Moreover, there was a good correlation between the Alternative Form and the Visual Reproduction subtest of the Wechsler Memory Scale – Revised, taking into account the test-retest reliability of the original version. Hence, this Alternative Form has the potential to be a useful addition to clinical practice with further refinement of the scoring criteria and development of normative data via administration to a larger sample that has a wider age range and intelligence.

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

INTRODUCTION

Leonardo Da Vinci stated "Whoever in discussion adduces authority uses not intellect but rather memory" (Knowles, 2004, p. 481). Indeed, there has been longstanding interest in memory function, but it has only been in the last 100 years that it has been scientifically evaluated. In fact, most knowledge about memory function has only been gathered from the second half of the 20th century. Memory dysfunction plays a critical role in many neurological and neuropsychological conditions, therefore its assessment becomes of major clinical value. This thesis pertains to an investigation of the development of an alternative form of the Visual Reproduction subtest of the Wechsler Memory Scale – Revised.

There have been various approaches to the scientific thinking of memory, but for a long time the assumption was held that memory function was a unitary process. Dissociable memory systems between short-term memory and long-term memory were postulated as early as 1890 by William James and again in 1949 by Donald Hebb, but it was not until the 1960's and early 1970's that the unitary concept of memory was challenged (Atkinson & Shiffrin, 1968; Tulving, 1972). Advances in experimental procedures, developments in neuro-imaging techniques and studies of brain damaged patients at an individual and group level have contributed to our understanding of the complexity of memory function.

1.1. Theoretical Models of Memory

Over the last four to five decades, a number of different models have been proposed to help better understand human memory and its component processes. Some of these models will be briefly presented below. More detailed descriptions are available in Broadbent, 1958; Waugh & Norman, 1965; Atkinson & Shiffrin, 1968; Craik & Lockhart, 1972; Tulving, 1972; Baddeley & Hitch, 1974; Cohen & Squire, 1980; Baddeley 1986, 1999, 2004; Squire, 1987, 1992; Tulving & Thomson, 1973; Squire & Butters, 1992; and Tulving & Schacter, 1994.

1.2. Models of Short-Term Memory

1.2.1. Information Processing Models

Many early models of memory were derived from theories of acquisition and were often based on the temporal duration of retention (Waugh & Norman, 1965; Atkinson & Shiffrin, 1968). One of the most influential models of its time was Atkinson and Shiffrin's (1968) 'modal' model of memory.

The modal model proposed memory as a sequence of three discrete processes in which information was transferred from one form of memory store to another. External stimuli from the environment initially entered a visual or auditory sensory memory system; namely iconic memory (visual sensory memory) and echoic memory (auditory sensory memory) for a period of a few seconds. Information was then selected via attentional processes and passed into short-term store. The short-term store was viewed to have a limited storage capacity, capable to hold information for approximately 30 seconds before it decayed. The assumption holds that the longer information was maintained in short-term store through sub-vocal rehearsal, the greater likelihood of its transfer into long-term memory. At each stage, information could be lost due to decay or interference from the presentation of new information. Long-term memory was proposed to have an unlimited storage capacity and was a relatively permanent store of information (Atkinson & Shiffrin, 1968; Baddeley 1994).

An alternative approach in the early 1970's was the 'levels of processing' model proposed by Craik and Lockhart (1972). The 'levels of processing' model was essentially concerned with the manner information was processed and the probability of its effective retrieval. That is, information could be processed in a variety of ways (i.e. via rehearsal of the information), but its likelihood of being stored in long-term memory was related to how 'deeply' or 'superficially' the information was encoded. The more meaning that was extracted from information, by relating it to previously acquired knowledge, the greater the probability it would be consolidated and stored in long-term memory (Craik & Lockhart, 1972; Baddeley, 1997).

The information processing models were primarily a unitary short-term storage system and proved to be oversimplified. Problems with the model came from evidence that holding an item in short-term store did not always guarantee its transfer into long-term memory (Baddeley, 1995; Parkin, 1997) and observations of brain damaged patients who had deficits in short-term memory, but relatively preserved learning ability (Shallice & Warrington, 1970).

In abandoning the assumption of a unitary system, Baddeley and Hitch (1974) proposed a more elaborate and flexible temporary store called working memory. The working memory system plays a greater role in the temporary maintenance and manipulation of complex cognitive processes such as comprehension, learning and reasoning (Emilien, Durlach, Antoniadis, Van der Linden & Maloteeaux, 2004).

1.2.2. Working Memory Model

In the original formulation by Baddeley and Hitch (1974), the working memory model consisted of two modality-specific storage systems, namely the articulatory loop (used for the processing of auditory-verbal information) and the visuo-spatial sketchpad (used for the processing of visual and spatial information). The articulatory loop (later revised to the phonological loop) and the visuo-spatial sketchpad are under the control of the central executive, which integrates, organizes and monitors the information from the two 'slave' systems (Baddeley, 1986; Baddeley, 1995).

According to more recent developments of the working memory model (Baddeley, 1999) the phonological loop is suggested to comprise two separate sub-components; a phonological store that maintains a limited amount of information in an acoustic code, and a sub-vocal rehearsal process that helps to maintain items in the phonological store. Auditory spoken information was assumed to gain immediate and necessary access to the phonological store.

There has been less research on the visuo-spatial sketchpad conducted than on the phonological loop, and consequently the components about the visuo-spatial sketchpad are less elaborated. There are suggestions of a subcomponent visuo-spatial passive store as well as a visual rehearsal process (Baddeley 1994; Cornoldi, & Vecchi, 2003). The visuo-spatial sketchpad has also been implied to play a role in the maintenance of visual representations of verbally encoded stimuli (Baddeley, 1999). Furthermore, the visuospatial sketchpad has recently been suggested to be more closely tied to the central executive than the phonological store (Miyake, Friedman, Shah, Rettinger & Hegarty, 2001).

In Baddeley's model, the central executive has a supervisory role in the integration and organization of information from the phonological loop and visuo-spatial sketchpad, and in the transfer and retrieval of information from long-term memory (Baddeley, 1986; 1997). The central executive has been described as an attentional control system responsible for strategy selection, control and execution of various processing tasks. The greater complexity of the central executive relative to the two 'slave' systems makes it the least well understood component of working memory and the most difficult to investigate (Baddeley 1999; 2004).

Criticisms of the concept of a central executive led Baddeley (2000) to propose an additional component to the model termed the 'episodic buffer.' The episodic buffer is assumed to have the capacity to temporarily store information from both slave systems and from long-term memory, to form an episodic representation available for conscious recollection (Baddeley, 2000; Gooding, Isaac & Mayes, 2005).

Despite ongoing challenges, research and refinement, the working memory model currently remains the most widely accepted conceptualisation of memory with considerable support from neuropsychological, experimental and neuro-imaging investigations (Baddeley, 2004; Miyake et al, 2001; Gooding et al, 2005).

1.3. A Model of Long-Term Memory

The organization of long-term memory has been the subject of ongoing debate and investigation. Over the last 30 years there has been considerable controversy regarding how long-term memory should be conceptualized, but it is now widely accepted that it is not a simple unitary system.

In early conceptualizations of long-term memory, Tulving (1972) proposed a dual classification system of semantic and episodic memory. Episodic memory referred to memory for personal events and experiences. Semantic memory represented the acquisition of knowledge about the world (e.g. the meaning of words, objects, concepts, facts and people) independent of the context in which this information was initially acquired (Tulving, 1972). Initially, the episodic and semantic memory systems were considered functionally distinct. This was supported by findings from studies of amnesic patients who showed impaired episodic learning, but relatively spared semantic memory (Kinsbourne & Wood, 1975; Parkin, 1982; Wood, Ebert & Kinsbourne, 1982).

Evidence against such a simple dissociation of long-term memory came from brain damaged patients with similarly impaired factual knowledge and loss of memory for personal experiences. Tulving (1983) revised his model and proposed that episodic memory was dependent on the integrity of semantic knowledge and had evolved from semantic memory. Episodic memory was assumed to share many properties and capabilities with the semantic system, but transcended it in the ability to recollect and relate experiences from the past in subjective time (Tulving, 2002; Schacter & Tulving, 1994).

Other investigators held the view that semantic memory was the result of accumulated episodic memories whereby the precise content was forgotten over time, such as memories from childhood (Cermak, 1984; Baddeley, 1994). Recent studies have described patients with impaired semantic knowledge yet relatively preserved episodic memory, as seen in semantic dementia, that refute Tulving’s hypothesis about the hierarchical nature of episodic and semantic memory (Graham, Simons, Pratt, Patterson

& Hodges 2000; Snowden, 2002; Emilien et al, 2004; Kazui, Hashimoto, Hirono & Mori, 2003; Mummery, Patterson, Wise, Price, Hodges & Vandenbergh, 1999; Hodges, Patterson, Oxbury & Funnell, 1992).

Some researchers have questioned the need to distinguish between episodic and semantic memory, subsuming them under a single declarative system. Squire (1992) proposed a declarative system of episodic and semantic memory, as both are available to conscious awareness and can be utilised in many novel situations. The relationship between episodic and semantic memory is an area of ongoing research, but the distinction nonetheless remains clinically useful, often for diagnostic issues.

A clear distinction between declarative forms of memory and procedural memory was proposed due to the finding that amnesic patients have an ability to show some forms of learning and not others (Cohen & Squire, 1980; Graf & Schacter, 1985). The relative preservation of procedural memory in amnesic patients means that the acquisition of skills can occur via an incremental learning process. This is believed to operate at an automatic level without conscious awareness (Papanicolaou, 2006). The concept of procedural memory was later revised and it became synonymous with the term non-declarative memory, thus encompassing a large set of heterogeneous non-conscious mnestic abilities, including the learning of skills and habits, non-associative learning, priming and simple classical conditioning (Squire, 1987; Squire, 1992).

Non-declarative forms of memory and learning are assumed to operate independently of the declarative system and are only indirectly accessible or demonstrable in circumstances similar to the original acquisition (Reber, Knowlton & Squire, 1996; Baddeley, 2004). A distinction between declarative memory (also called explicit memory) and non-declarative memory (or implicit memory) systems is currently the most widely accepted model of the organization of long-term memory (Squire, 1987, 1992; Graf & Schacter, 1985; Baddeley, 2004; Schacter, Chiu & Ochsner, 1993).

1.4. Memory Processes

The information processing models delineate encoding, storage and retrieval memory processes (Craik & Lockhart, 1972; Tulving & Thompson, 1973). Encoding reflects the processing of incoming information and the laying down of a memory trace involving two phases; an acquisition and consolidation phase. The acquisition phase registers external inputs into sensory buffers and sensory analysis stages, while the consolidation phase creates a stronger representation of this information over time (Nadel & Moscovitch, 1997). The consolidation process converts information from short-term memory into long-term memory and is susceptible to modification (Emilien et al, 2004). Storage is the result of the acquisition and consolidation phases, where a permanent record of information is maintained. The retrieval process re-constructs stored information from long-term memory to create a conscious representation or to execute a procedural learnt behaviour or motor response, such as riding a bicycle (Lezak, Howieson & Loring, 2004).

There is substantial evidence that memory processes are organised in a dynamic way and continually being modified and re-organised (Kopelman, 2002; Nadel, 1992; Reber, et al, 1996). Investigators in memory research have accepted that no model can entirely account for the complex and multivariate nature of the processes and organization of memory (Nade, 1994; Schacter & Tulving, 1994).

1.5. The Neuroanatomy of Memory

Over the years, cumulative and systematic research utilizing animal models, observation, neuropathological findings of localised brain injuries, ablation studies, and more recently from the use of sophisticated neuro-imaging analyses has assisted in the understanding of particular brain structures important for effective memory function (Scoville & Milner, 1957; Penfield & Perot, 1963; Mishkin, 1978; Zola-Morgan, Squire & Amaral, 1986; Mayes & Montaldi, 2001).

The structures implicated in sub-serving a crucial role for declarative memory are the hippocampus and related anatomical structures of the medial temporal lobe, the medial diencephalic structures and the basal forebrain (Zola-Morgan & Squire, 1993; Squire, Knowlton & Musen, 1993).

1.5.1. The Medial Temporal Lobes

Much of the understanding of the key role of the medial temporal lobes in memory function stems from Scoville and Milner's (1957) description of patient H.M who underwent bilateral resection of the medial temporal lobes for the relief of intractable epilepsy. Patient H.M developed severe anterograde amnesia (a severe impairment in new learning ability) and retrograde amnesia (impaired recollection for events prior to the surgery) that extended several years. This amnesia persisted for the rest of his life. Animal research in monkeys involving similar lesions approximating the damage sustained by patient H.M reproduced similar features of memory impairment in that species (Zola-Morgan & Squire, 1993). The evidence from H.M and animal models strongly implicated the hippocampal complex (including the dentate gyrus and subiculum) as critical structures for the acquisition of new episodic and semantic memories (Mishkin, 1982; Tulving & Markowitsch, 1998; Broadbent, Clark, Zola & Squire, 2002; Deweer, Pillon, Pochon & Dubois, 2001; Lee, Yip & Jones-Gotman, 2002).

Animal research and human studies suggest that the surrounding cortical areas, with a close anatomical relationship to the hippocampus, are also crucial for memory function (Zola-Morgan & Squire, 1985; Emilien et al, 2004; Rosenbaum, Kohler, Schacter, Moscovitch, Westmacott, Black, Gao & Tulving, 2005). In particular, the entorhinal, perirhinal and parahippocampal cortices are implicated as important for declarative memory (Squire & Shimamura, 1996; Zola-Morgan, Squire, Amaral & Suzuki, 1989; Miller, Lai & Munoz, 1998). Evidence has shown that the perirhinal and the parahippocampal structures provide nearly two thirds of the cortical input to the entorhinal cortex (Tulving & Markowitsch, 1998). As the entorhinal cortex is a major source of projections to the hippocampus and the dentate gyrus, the anterograde amnesia becomes more severe when these cortical regions are damaged as well (Zola-Morgan & Squire, 1993; Zola, 1998; Zola-Morgan et al, 1989; Graham et al, 2000; Nadel, 1994).

The neuroanatomical damage typically associated with anterograde amnesia from virus induced lesions, such as herpes simplex encephalitis, involves the medial temporal lobe structures (including the hippocampus, entorhinal, perirhinal and parahippocampal cortices), polar limbic regions and the amygdala (Zola-Morgan & Squire, 1993; Rotenberg & Weinberg, 1999; Squire & Zola-Morgan, 1991). Furthermore, greater severity of memory deficits has been described with lesions involving the amygdala, hippocampus, and the perirhinal and parahippocampal cortices when compared with hippocampal damage alone (Zola-Morgan, et al, 1989; Gleisnner, Helmstaedter, Schramm & Elger, 2002; Miller, Lai & Munoz, 1998).

Even partial damage to the hippocampus has been found to produce substantial memory impairment in humans and monkeys (Zola-Morgan & Squire, 1993; Mishkin, Vargha-Khadem & Gadian, 1998). Findings from patient R.B, who had circumscribed memory impairment as a result of a hypoxic brain injury, revealed that a bilateral lesion involving the CA1 field of the hippocampus was sufficient to produce severe anterograde amnesia (Zola-Morgan, Squire & Amaral, 1986). Magnetic Resonance Imaging studies also indicate a reduction in hippocampal size in those with clinically significant memory impairments (Squire et al, 1990; Markowitsch, 2003; Mayes & Montaldi, 1999).

There is less agreement concerning the role of the hippocampal complex in the retrieval of information and retrograde amnesia (Emilien et al, 2004). However, lesions that extend to the entorhinal and perirhinal cortices in association with the hippocampus are implicated in a severe retrograde amnesia extending several decades (Rempel-Clower, Zola, Squire & Amaral, 1996; Baxter & Murray, 2001; Corkin, Amaral, Gonzalez, Johnson & Hyman, 1997; Zola-Morgan et al, 1986; Markowitsch, 2000).

Recent data from studies of humans, monkeys and rats suggest the hippocampus plays an essential role in novelty discrimination (Broadbent et al, 2002; Squire & Zola, 1998; Dolan & Fletcher, 1997; Redoblado, Garyson & Miller, 2003). Furthermore, functional brain imaging studies have shown selective activation of the right mesial temporal region in response to novel and unfamiliar pictures and words (Wiggs, Weisberg, & Martin, 1999; Iidaka, Sadato, Yamada & Yonekura, 2000; Mayes & Montaldi, 2001).

Studies of amnesic patients have consistently suggested that the hippocampus plays a key role in spatial memory (Scoville & Milner, 1957; Nadel, 1992; Corkin, et al, 1997; Emilien et al, 2004; Nunn, Polkey & Morris, 1998; Winocur, Moscovitch, Caruana & Binns, 2005; Moye, 1997). Specifically, the parahippocampal cortex has been suggested to be especially important for spatial memory and the perirhinal cortex as important for visual memory (Markowitsch, 2000; Baxter & Murray, 2001), suggesting that the hippocampus and surrounding structures are essential for the consolidation of information into memory irrespective of modality.

Hemispheric specialisation of the left and right medial temporal structures for verbal declarative memory and non-verbal learning are widely documented in the literature (Parkin, 1997; Frisk & Milner, 1990; Lee et al, 2002). Damage to the left temporal lobe has consistently demonstrated deficits in the learning of verbal information. Although findings are not as consistent, damage to the right temporal lobe has been shown to cause non-verbal memory deficits (Chelune & Bornstein, 1988; Squire & Butters, 1992; Baxendale, 1997; Squire, 1986; Jones-Gotman, 1986; Chelune & Bornstein, 1988; Naugle, Chelune, Cheek, Luders & Awad, 1993). The inconsistencies in the findings with regard to non-verbal memory are considered to reflect methodological issues and the difficulty in developing a valid measure of non-verbal memory. This issue will be discussed later in section 1.12.

The specific role of the amygdala in memory has been the subject of much research. Animal models and human studies have identified the amygdala as significant in the encoding of emotionally arousing stimuli and memory (Markowitsch, 2000; Miller et al, 1998; LaBar & LeDoux, 2003). The amygdala has been implicated to play a potentially central role in the development of conditioned fear and influence the behavioural response to a neutral stimulus from previous experience (LeDoux, 1992). Furthermore, perceived stressful or dangerous stimuli are suggested to foster highly affective sensory representations that can disrupt the consolidation functions of the hippocampus from traumatic episodes (Zola-Morgan & Squire, 1993; LeDoux, 1992; Nadel & Jacobs, 1998).

1.5.2. The Diencephalon

Early investigations of memory impairment involving the medial diencephalon came from the study of patients with Wernicke Korsakoff Syndrome (Butters et al, 1995). Damage to the dorso-medial thalamic nucleus, mamillary bodies, mamillothalamic tract, fornix and areas adjacent to the third ventricle were implicated in extensive anterograde and retrograde deficits characteristic of Wernicke Korsakoff Syndrome and other diencephalic amnesias (Parkin, 1997; Victor, Adams & Collins, 1989; Butters, Salmon, Cullum, Cairns, Troster, Jacobs et al, 1988; Carr, 1982; Gaffan, Parker & Easton, 2001). Significant memory difficulties because of damage specific to the internal medullary lamina of the thalamus have been documented in recent studies with experimental animals and amnesic patients (Van der Werf, Witter, Uylings & Jolles, 2000; Zola-Morgan & Squire, 1993). The severe amnesia that arises from thalamic strokes involving bilateral infarction of medial thalamic structures has been well documented (Markowitsch, 2000; Aggleton, & Brown, 1999). Retrograde amnesia has been associated with lesions to the nuclei in the mamillary bodies and to the thalamus and to its interconnecting pathways to cortical regions, particularly the forebrain (Van der Werf, Scheltens, Lindeboom, Witter, Uylings & Jolles, 2003).

1.5.3. The Frontal Lobes

Lesions restricted to the frontal lobes ordinarily do not produce a severe amnesia of the kind commonly associated with damage to the medial temporal lobe and diencephalon, but impairment can occur for related aspects of memory (Wirsen & Ingvar, 1991; Dolan

& Fletcher, 1997; Buckner, Kelley & Petersen, 1999; Busch, Booth, McBride, Vanderploeg, Curtiss & Duchnick, 2005). Damage to the frontal lobes can result in impairment to prospective memory (the ability to remember an upcoming event or scheduled commitment), meta-memory (knowledge about memory), memory for temporal order (sequencing facts from the past in order of their occurrence), context and source memory (how specific knowledge was acquired), working memory and retrieval processes (Parkin, 1997; Baddeley, 2004; Zola, 1998; Zola-Morgan & Squire, 1993; Nadel & Moscovitch, 1997; Fletcher, Shallice, Frith, Frackowiak & Dolan, 1998; McDaniel & Einstein, 2007).

Functional neuro-imaging studies have identified bilateral prefrontal involvement during encoding and retrieval of episodic memories (Nyberg, McIntosh & Tulving, 1998; Lee, Robbins, Pickard & Owen, 2000; Fletcher, Shallice & Dolan, 1998; Fletcher et al, 1998). Both the encoding and retrieval of verbal material has been shown to activate the left prefrontal cortex. The left prefrontal cortex has been associated with active and strategic operation of mnemonic representations (Iidaka et al, 2000; Klingberg & Roland, 1998) and the success of a retrieval attempt. The right prefrontal cortex has been associated with monitoring processes and retrieval attempt, rather than retrieval success (Iidaka et al, 2000; Wiggs et al, 1999).

There are still arguments with respect to the lateralization of prefrontal activation during encoding of non-verbal material (Nyberg et al, 1998). In some studies, pattern specific encoding predominately activated the right prefrontal cortex. Other studies have observed left prefrontal activation during encoding of non-verbal material such as object location and faces (Klingber & Roland, 1998; Lee et al, 2000). Left prefrontal activation in the encoding of non-verbal information has been attributed to the use of verbal strategies and semantic representations of visual stimuli (Iidaka et al, 2000).

The basal forebrain (including the medial septal nuclei, the diagonal band of Broca, nucleus basalis and substantia innominata) are also implicated in normal memory function. The basal forebrain nuclei are a primary source of acetylcholine innervations to the medial temporal lobe, particularly the hippocampus, and to other areas of the cerebral cortex (Zola-Morgan & Squire, 1993). Decreased activity of cholineacetyltransferase in the hippocampus and in the cortex as well as reduced cell numbers in the basal forebrain are evident in patients with Alzheimer’s disease, who exhibit memory impairment early in the neurodegenerative process (O’Connor & Verfaellie, 2004). The memory impairments from anterior communicating artery aneurysms are less severe than the other amnesias but are relatively consistent with frontal lobe dysfunction. The disruption of frontal circuits to medial temporal and diencephalic structures and systems have also been implicated (Parkin & Leng, 1993).

1.5.4. From Brain Structures to Brain Systems

The connections between neural structures are suggested to play as an important role in normal memory functioning as the structures themselves. Rather than there being a single site for memories, a number of contributions from many cortical and subcortical structures within brain systems are considered to sub-serve functional roles for memory performance (Nadel & Moscovitch, 1997; Iidaka et al, 2000).

The interconnections between medial temporal lobes, the diencephalon and the frontal lobes have been identified as essential for the acquisition and consolidation of declarative memory. The complexity of the circuitry extends to include the neocortex, which is suggested as the repository for the storage of long-term memories, gradually supporting the storage of information independent of the frontal, medial temporal and diencephalic systems (Zola-Morgan & Squire, 1993; Nadel & Moscovitch, 1997; Mayes & Montaldi, 1999; Zola-Morgan & Squire, 1990).

The evidence for multiple memory systems from animal and individual case studies suggests distinct neural systems for impairments of declarative learning and memory. Mishkin (1982) proposed a model that consisted of two distinct memory pathways that needed to be lesioned in order to produce severe impairments in memory function, namely the hippocampal and amygdalar pathways. The hippocampal pathway projects from the hippocampus to the mamillary bodies via the fornix that travels along the mamillothalamic tract to the anterior nucleus of the thalamus, possibly on to the cingulate cortex and back to the hippocampus. The amygdalar pathway projects from the amygdala to the dorsomedial thalamic nucleus on to the orbito-frontal cortex before returning to the amygdala (Mayes, 1986; Markowitsch, 2003). This dual circuit model has gained support, but it is argued that damage to the hippocampal system alone may be sufficient to produce a severe amnesic syndrome (Parkin & Leng, 1993; Papanicolaou, 2006).

Non-declarative memory represents relatively independent memory systems mostly involving anatomical structures outside the medial temporal lobes and diencephalon (Squire, Knowlton & Musen, 1993). However, the corticostriatal systems involving reciprocal connections between the neocortex and the basal ganglia are seen as subserving non-declarative memories (Iidaka et al, 2000). The type of information processing required in performing a specific task typically determines the brain regions that are involved (e.g., the pre-motor cortex is often implicated in motor tasks such as walking) (Markowitsch, 2003).

Historically, investigations in the neuro-anatomical bases of non-declarative memory systems have been limited. The growing research on the nature of implicit memory has been fuelled by interest in preserved skills of individuals with severe memory dysfunction (Zola-Morgan & Squire, 1993; Lezak et al, 2004; Papanicolaou, 2006).

1.6. Memory Dysfunction

Memory dysfunction can be found in diverse neurological disorders including epilepsy, aneurysms, brain tumours, cerebrovascular accidents, encephalitis, anoxia, multiple sclerosis, traumatic brain injury, chronic alcohol abuse and degenerative diseases of the brain (such as Alzheimer’s disease and fronto-temporal dementia) (Chelune & Bornstein, 1988; Squire, 1986; Kapur, 1988; Baddeley, 1995; Squire & Shimamura, 1996; Tulving, 2002). Deficits in memory often manifest themselves as initial symptoms in certain disease processes (e.g., dementia) before deficits in other areas of functioning present themselves (Roid, Prifitera & Ledbetter 1988; Papanicolaou, 2006).

The most severe disorders of declarative memory have typical features that are collectively described as the amnesic syndrome. Typically, the amnesic syndrome is defined as a circumscribed deficit in memory in the context of relatively preserved intellectual function, insight, language and social skills (Scoville & Milner, 1957; Parkin & Leng, 1993; Papanicolaou, 2006). Severe anterograde amnesia is a cardinal feature of the amnesic syndrome, but it can also occur with other memory syndromes (such as dementia). The severity of the retrograde amnesia can be variable, depending on the extent and nature of structural compromise and can occur in the absence of anterograde memory loss, as in cases of focal retrograde amnesia (Kapur, 1993). Moreover, the retrograde amnesia typically follows a temporal gradient. That is, memories from the distant past (such as childhood) are better preserved compared to recent memories (Parkin, 1997).

Despite the severity and extent of anterograde memory impairment, immediate memory and working memory capacity (the ability to hold and manipulate a limited amount of information in mind for a very brief period of time) are usually preserved in amnesic patients (Papanicolaou, 2006). Moreover, well established semantic memories also remain relatively unimpaired, as evidenced by intact knowledge of tasks and word meanings (Baddeley, 1997). Non-declarative memory, as measured by perceptuo-motor skill acquisition, priming, and classical conditioning are often preserved, since these forms of memory are independent of the medial temporal structures, which is the site of damage in many amnesic disorders (Baddeley, 1995; Squire, 1986; Parkin, 1997; O'Connor & Verfaellie, 2004).

Degenerative diseases can produce cognitive changes similar to those of the amnesic syndrome, congruent with specific areas of pathology (Lezak et al, 2004; Kopelman, 2002). Temporal, diencephalic and frontal regions are readily implicated in degenerative diseases such as Alzheimer's disease, frontotemporal dementia, and multi-infarct dementia (Squire, 1987; Squire, 1992; Zola-Morgan & Squire, 1993). The pattern of deficits in Alzheimer's disease reflects an impairment of episodic memory, a retrograde amnesia with a temporal gradient, and the compromise of at least one other cognitive function (Kopelman, 2002). Neuro-degenerative diseases such as Huntington's disease, Multiple Sclerosis, Parkinson's disease and Progressive Supranuclear Palsy are collectively known as subcortical dementias. The memory impairment in patients with subcortical dementias is more pronounced for procedural learning tasks than for tasks involving declarative memory (Howieson & Lezak, 2004).

Generally, memory loss in dementia differs from amnesia as immediate memory span and semantic memory is often affected in dementia. The progressive nature of dementia results in increasing disorientation, personality and behavioural change, intellectual and cognitive compromise, as opposed to the static nature of an amnesic syndrome (Walsh & Darby, 2001; Parkin & Leng, 1993; Papanicolaou, 2006).

The cognitive profile of individuals with cognitive dysfunction and certainly the pattern of memory difficulties needs to take into account not only aetiological factors, but a range of potential metabolic, medical, and psychological problems (O'Connor & Verfaellie, 2004). A comprehensive neuropsychological evaluation that considers all these factors is important in order to determine which functions are compromised and which are preserved.

1.7. Assessment of Memory

The evaluation of memory can facilitate identification of a deficit in an important cognitive function, diagnosis of an underlying disorder or cause of memory problems, measurement of the extent and severity of dysfunction, contribute to treatment and management, help determine whether a deficit is organic or functional in origin, assess changes in functioning over time, and help evaluate the success of rehabilitation interventions (Mayes, 1986; Eslinger, 2002; Gfeller, Meldrum & Jacobi, 1995; Squire, 1986; Lezak et al, 2004). The assessment of memory for research purposes is important for establishing the neuropsychological profiles of particular clinical populations with different neuro-pathological conditions, and also for evaluating theoretical conceptualizations of memory (Howieson & Lezak, 2004; Wilson, 2004).

The development of neuropsychological measures often reflects the theoretical conceptualization of memory at the time of its development. Traditionally, memory assessment and test development have largely focused on declarative memory and so this area will be the focus of this review.

Standardized measures of memory need to meet certain criteria if they are to be clinically useful. Memory tests need to deliver reliable and consistent results and they need to measure the construct that they purport to measure (Cohen, 1988; Tabachnick & Fidell, 1996). A measurement procedure is considered reliable if relatively stable and consistent scores are produced when an individual is tested under the same conditions on another occasion (Gravetter & Wallnau, 1996). The reliability of a measure can be evaluated in a number of ways; including test-retest procedures, split-half reliability, Cronbach's alpha, and standard errors of measurement. Although a measure that is not reliable cannot be valid, a measure may not be valid even if it is reliable (Cohen, 1988; Tabachnick & Fidell, 1996).

The validity of a measure is an evaluative judgment of the adequacy and appropriateness of inferences to be made based on test performance and the potential of the score to reflect what the test purports to measure (Gravetter & Wallnau, 1996). The validity of a measure is generally considered the most important psychometric property of an instrument and must be established so that erroneous conclusions are not made regarding memory functioning (Tabachnick & Fidell, 1996; Cohen, 1988). The construct validity (the verification of the measure and the theory of the construct it is intended to measure), content validity (how well inferences can be drawn from one item to another item of a measure), criterion-related validity (relating one measure to another measure of the same construct), convergent validity (how well a measure is associated with a measure it should correlate with) and discriminant validity (how poorly a measure correlates with a measure it should poorly be associated with) of a measure can be determined (Gravetter & Wallnau, 1996; Willmes, 2003).

For measurement of memory function to be useful, there must be comprehensive statistical and normative data available for the measures used so that abnormalities can be clearly identified and the severity of deficit quantified (Lezak et al, 2004). Scoring systems also need to be reliable and test ceiling and floor effects need to be minimised to ensure that all levels of memory performance can be quantified. That is, the measure is not too easy for some individuals and too difficult for others (Mayes, 1995; Eadie & Shum, 1995). The ease and duration of administration of a measure and the availability of alternative forms are considered to be important factors in establishing the clinical utility of a measure (Willmes, 2003; Wilson, 2004).

The inclusion of testing procedures that allow for inferences regarding the relative integrity of encoding, storage and retrieval processes are currently viewed as essential in the comprehensive evaluation of memory functions (Gass, 1995). If a psychometric measure incorporates both immediate and long-term memory processes, the contribution of each process needs to be established. The assessment of memory also needs to quantify the extent that competing information interferes with the retrieval of old information and new learning ability (Butters, Delis & Lucas, 1995; Lezak et al, 2004; Emilien et al, 2004). In evaluating the basis of retrieval deficits, a recognition procedure is often the most direct method of assessment (Gass, 1995). The administration of a cued recall procedure may facilitate the recall of information from long-term memory and hence provide important information regarding the extent of retrieval difficulties.

The assessment of declarative memory has been further refined to include tasks that reflect the distinction of verbal and non-verbal memory abilities (Smith et al, 1992). The assessment of auditory-verbal memory at different levels of complexity typically involves tests of immediate memory span, list learning, story memory and paired associate learning tasks. Immediate memory measures such as digit span (Wechsler, 1987; 1997) are serial recall tasks that require the repetition of digits in a sequence. In addition, tests of increasing length of sentences have been developed as measures of immediate memory capacity, as deficits on these tasks are common in many disorders of memory (Goodglass & Kaplan, 1983).

Memory for lists of words include measures such as the Rey Auditory Verbal Learning Test (Rey, 1964b), the California Verbal Learning Test (Delis, Kramer, Kaplan & Ober, 1987) and its second revision (Delis, Kramer Kaplan & Ober 2000), the Hopkins Verbal Learning Test (Brandt, 1991) and the Selective Reminding Test (Hannay & Levin, 1985). These tests often assess immediate memory, rate of learning and learning ability, the presence of interference effects, long-term memory and recognition. Verbal paired associate tasks involve the learning of word pairs and the subsequent recall of word associations that evaluate the integrity of the left hippocampal system (Wechsler, 1945; Wechsler, 1987).

Tests of story memory are usually developed to simulate language used in everyday conversation and typically evaluate memory for meaningful information. The most widely used test of story memory is the Logical Memory subtest from the Wechsler Memory Scale and its revisions (Wechsler, 1945; Wechsler 1987; Wechsler, 1997) although several other story memory measures have been developed (Heaton, Grant & Williams, 1991; Lezak, 1995; Williams, 1991).

The assessment of non-verbal memory typically involves tests of immediate visual span, memory for visually presented stimuli (such as designs and faces) and recognition memory measures (Moye, 1997; Lezak et al, 2004). Immediate visual measures such as the Corsi blocks (Milner, 1971) and the spatial span subtest (Wechsler, 1981; Wechsler, 1997) have been developed. They involve the repetition of a particular sequence of steps in the same manner as the sequence was presented to evaluate immediate memory for visuo-spatial information.

The most common memory tests for visually presented stimuli require the reproduction of geometric designs and include the Rey Complex Figure Test (Rey, 1964; Meyers & Meyers, 1995), Biber Figure Learning Test (Glosser, Goodglass & Biber, 1989), Benton Visual Retention Test (Benton, 1992), and the Visual Reproduction subtest of the Wechsler Memory Scale and its revisions (Wechsler, 1945; Wechsler 1987; Wechsler 1997). These tests are designed to assess immediate and long-term memory for visuo-spatial material. However, other neuropsychological deficits such as visuo-perceptual problems and drawing difficulties can compromise performance and lead to misdiagnosis of the extent and nature of non-verbal memory impairments (Smith et al, 1992; Larrabee & Curtiss, 1995; Haut, Weber, Wilheim, Keefover & Rankin, 1994; Loring, 1989; Gfeller et al, 1995; Moye, 1997; Bowden, Ritter, Carstairs, Shores, Pead, Greeley et al, 2001). A limited number of standardised measures have included recognition procedures to provide information on the relative contributions of encoding, storage and retrieval on visual memory performance (Glosser et al, 1989; Meyers & Meyers, 1995).

Recognition measures have been developed to help overcome problems with tasks that require a motor response, but these are criticised for not having immediate and delayed free recall components. Tests of recognition memory include the Continuous Recognition Memory Test (Hannay, Levin & Grossman, 1979), Figural Memory subtest (Wechsler, 1987), Faces and Family Pictures subtests (Wechsler, 1997) and the Warrington Recognition Memory Test (Warrington, 1984).

In clinical practice, standardized test batteries are typically administered so a variety of memory abilities can be evaluated with the advantage of the same normative sample. One of the most prominent and frequently used batteries in the clinical assessment of memory has been the Wechsler Memory Scale and its revisions. The Wechsler Memory Scale (1945) has undergone two revisions since its inception: the Wechsler Memory Scale – Revised (1987), and the Wechsler Memory Scale – Third Edition (1997).

1.8. The Wechsler Memory Scale

The Wechsler Memory Scale has arguably been one of the most widely used measures of memory. The Wechsler Memory Scale was developed due to the lack of a ‘rapid, simple and practical’ standardised measure in the clinical assessment of memory and its disorders (Wechsler, 1945, p. 3).

1.8.1. Structure and Content of the Wechsler Memory Scale

The Wechsler Memory Scale consisted of seven subtests: (1) Personal and Current Information (e.g., questions such as name, age, date of birth, current government leaders); (2) Orientation (e.g., questions relating to place and time); (3) Mental Control (e.g., counting backwards, reciting the alphabet and counting by 3’s under time constraints); (4) Logical Memory, which is a test of immediate auditory memory (e.g., reciting and recalling the content presented in two stories); (5) Memory Span/Digits Span, a test of attention and immediate auditory memory (e.g., repeating digits and reciting digits in reverse order); (6) Visual Reproduction, which is a test of immediate visual memory (e.g., the drawing of geometric designs from memory); and (7) Associate Learning (e.g., verbal recall of the correct response to six semantically related (easy) and four unrelated word pairs (hard) over three trials) (Wechsler, 1945; Mitrushina, Boone & D’Elia, 1999).

An alternative form (Form II) of the Wechsler Memory Scale was also developed for the interchangeable use in repeated memory assessments. Form II was comparatively matched to the design of form I, consisting of seven subtests, but with changes made to the content of some subtests (i.e., Mental Control, Logical Memory, Visual Reproduction and Associate Learning).

1.8.2. Scoring the Wechsler Memory Scale

The scoring of each subtest was based on a system awarding one point for a correct response and zero points for an incorrect response. The raw scores of each of the seven subtests were summed and an age correction factor was added to yield a single memory score, the global Memory Quotient (Wechsler, 1945). The Memory Quotient was intended to be comparable with performance on measures of intellectual functioning.

Administration procedures, such as test instructions and materials were relatively straightforward, but problems with the scoring procedures existed. The scoring procedures for the Logical Memory and Visual Reproduction subtests were brief, imprecise and inadequate and resulted in low inter-scorer agreement (Prigatano, 1978; Loring & Papanicolaou, 1987; Herman, 1988).

A problem with the concept of the Memory Quotient was also identified. The use of a single score to reflect memory functioning assumed that memory was a unitary construct. A unitary view of memory functioning was inconsistent with clinical and research findings and is a theoretical concept contemporarily rejected (Lezak et al, 2004). The use of a single score was further criticised for its potential to mask variability among subtest performances as it did not allow the differentiation of distinct memory functions, consequently limiting its sensitivity as a diagnostic instrument (Parkin & Leng, 1993).

Questions over the clinical utility of the IQ – Memory Quotient discrepancy in identifying true memory impairment and the usefulness of the Wechsler Memory Scale in differentiating organic from psychiatric memory problems were raised (Prigatano, 1978). For example, individuals with amnesic syndromes performed well on the measure mainly because it did not test long-term memory. The high correlation between the Memory Quotient score and intelligence measures questioned whether the Wechsler Memory Scale was a measure of memory or intellectual functioning (Erickson & Scott, 1977).

1.8.3. Reliability of the Wechsler Memory Scale

The Wechsler Memory Scale manual provided no empirical data to support the claim that it was a reliable measure (Prigatano, 1978). There was no information regarding the test-retest reliability of the Wechsler Memory Scale or the internal consistency of the subtests of either Form I or Form II (Prigatano, 1978). A significant discrepancy in the scoring of the Logical Memory and Visual Reproduction subtests suggested poor inter-rater reliability (Mitchell, 1987).

Although an approximate equal level of difficulty between the two forms was documented, the Visual Reproduction subtest has been found to be easier on Form II and the Associate Learning subtest easier on Form I, making interpretation of equivalence between the two forms somewhat tenuous (Prigatano, 1978; Ivison, 1993).

1.8.4. Normative Data of the Wechsler Memory Scale

Wechsler's reports of the standardization of the original Wechsler Memory Scale are inconsistent. He noted that the standardization of Form I was based on 200 normal participants aged 25-50 years. However, data in the manual was only shown for two groups of individuals aged 20-29 years (n = 50) and aged 40-49 years (n = 46). No data on the standardization of Form II was given and norms were only available for Form I (Mitrushina et al, 1999).

The Wechsler Memory Scale data published has been extensively criticised for inadequate provision of norms across age groups, a small and restricted sample, and limited information regarding characteristics of the standardization sample (Loring & Papanicolaou, 1987). The percentage of men and women in the sample were not reported nor were differences in performances on subtests between the sexes. The restricted age range of the normative sample limited the clinical utility of the measure, as lower memory performance is seen in normal elderly populations.

In an attempt to address some of the standardization problems of the Wechsler Memory Scale, normative studies have provided some initial norms for adolescents and older adults of various population groups (Ivison, 1977; Cauthen, 1977; Prigatano, 1978). For a comprehensive review and critique of the normative studies undertaken for the Wechsler Memory Scale, see Mitrushina et al 1999; Mitrushina, Boone, Razani & D'Elia, 2005.

1.8.5. Validity of the Wechsler Memory Scale

Numerous factor analyses have appeared in the literature revealing a three factor structure of the Wechlser Memory Scale of 1) immediate learning and recall 2) attention and concentration 3) orientation (Mitrushina et al, 1999; Skilbeck & Woods, 1980).

The construct validity of the Wechsler Memory Scale as primarily a measure of short-term verbal memory has been consistently supported. Research demonstrated considerable face validity of the Wechsler Memory Scale as primarily a measure of verbal memory with six of the seven subtests being verbal in nature (Chelune & Bornstein, 1988).

The Wechsler Memory Scale has been found to be sensitive to memory disturbances associated with left temporal lesions, but generally insensitive to the detection of memory dysfunction related to right temporal damage (Prigatano, 1978; Larrabee, Kane & Shuck, 1983; Skilbeck & Woods, 1980). The insensitivity of the Wechsler Memory Scale to right hemispheric dysfunction was suggested to be the result of the inclusion of only the Visual Reproduction subtest as a measure of visual memory. In addition, the predominant verbal nature of the Wechsler Memory Scale implied that individuals who have verbal expressive problems could also be unduly penalised (Erickson & Scott, 1977; Prigatano, 1978).

The composition of the Wechsler Memory Scale was further challenged because of its inclusion of measures thought to reflect other cognitive functions, like attention and concentration (e.g., Information, Orientation, Mental Control and Digit Span subtests) (Erikson & Scott, 1977; Prigatano, 1978). The inclusion of tests that addressed other functions had the potential to overestimate or mask important patterns of memory performance (Delis, 1989).

In addition, the lack of measures for evaluating the retention of learned material over time and to assess recognition were a significant shortcoming of the Wechsler Memory Scale, as these measures are now well established to provide valuable clinical information (Mayes, 1995). In an attempt to address some of the limitations of the original scale, two independent variations of the Wechsler Memory Scale were developed.

Russell (1975) developed delayed recall procedures for the Logical Memory, Visual Reproduction and Paired Associate task, improved the scoring procedures for the Logical Memory subtest and introduced the calculation of saving scores (i.e. percent retention scores) for both the Logical Memory subtest and the Visual Reproduction subtest. The additional procedures considerably improved the utility of the Wechsler Memory Scale (Larrabee, Kane, Schuck & Francis, 1985). However, a problem with Russell's revision was the normative data incorporated cuing procedures, making it difficult to compare to the original scale. In addition, no data was provided on how individuals of different levels of intellectual ability performed on these procedures (Waldmann & Dickson, 1991). Nonetheless, the Russell revision was a popular and widely used variant in clinical settings.

The Boston Revision of the Wechsler Memory Scale (Milberg, Hebben & Kaplan, 1986) included delayed recall and recognition measures for the Visual Reproduction, Logical Memory and Paired Associate subtests. Additional procedures included copy and perceptual match trials to cater for sensorimotor deficits. The inclusion of recognition and copy trails immediately after an initial presentation was a major criticism as it allowed for multiple exposures to the material to confound performance. Therefore, the Boston revision was not comparable to the original Wechsler Memory Scale. In addition, no detailed normative data was made available for this revision.

Further improvements of the Wechsler Memory Scale included the development of six parallel forms for the Associate Learning subtest (Nott, 1975) and a Visual Associate Learning task as a comparative measure to the Verbal Associate Learning subtest (Fowler, 1969).

The Wechsler Memory Scale was an early step in the development of clinical assessment tools. The Logical Memory subtest and the hard pairs on the Associate Learning task were considered good predictors of an amnesic syndrome and subsequently used as screening measures (Parkin & Leng 1993, Lezak, 1995). Despite its shortcoming, the relatively quick administration time of approximately 30 minutes made the Wechsler Memory Scale a popular test amongst many clinicians for a number of decades. It was not until 1987 that a review of the Wechsler Memory Scale was published.

1.9. Wechsler Memory Scale – Revised

The Wechsler Memory Scale – Revised was revised to address the limitations associated with the original scale. The Wechsler Memory Scale – Revised was described as a ‘diagnostic and screening device’ to better assist in the clinical evaluation of memory functions (Wechsler, 1987, p. 1).

1.9.1. Structure and Content of Wechsler Memory Scale – Revised

The Wechsler Memory Scale – Revised retained six of the subtests from the original Wechsler Memory Scale. With the inclusion of three new visual memory measures, the Wechsler Memory Scale – Revised comprised nine subtests in total: (1) Information and Orientation; (2) Mental Control; (3) Figural Memory; (4) Logical Memory; (5) Associate Learning (now called Verbal Paired Associates); (6) Visual Paired Associates; (7) Visual Reproduction; (8) Digit Span; and (9) Visual Memory Span.

The changes incorporated in the structure and content of the Wechsler Memory Scale – Revised included the combining of the Information and Orientation subtests as screening measures that no longer contributed to a memory score and the elimination of extra credit for fast performances on the Mental Control subtest. The inclusion of additional trials of shorter digit sequences for both forward and backward series on the Digit Span subtest was also part of the revision. The Logical Memory subtest retained the first story with only some minor amendments made to the content. A new story was developed to replace the second story in an attempt to make them more equivalent in difficulty.

New designs were developed that replaced the last two of the four designs of the original Visual Reproduction subtest. The Verbal Paired Associates subtest maintained most of the original word pairs, with the elimination of two easily learned word pairs to shorten the subtest. An additional three trials of the Verbal Paired Associate subtest were included so that the learning of the word pairs could be examined over a longer time frame, but these extra trials did not contribute to the total score. Delayed recall procedures were also developed for each of the Logical Memory, Visual Reproduction and Verbal Paired Associate measures in response to the criticisms of the original Wechsler Memory Scale.

In an effort to provide a more balanced assessment of visual memory the Figural Memory and Visual Paired Associate subtests were developed. Also a visual attention task, Visual Memory Span was developed as the analogue of the Digit Span subtest (Loring, 1989).

The Figural Memory subtest was devised to measure memory for figural stimuli. The task involved the identification of target abstract visual patterns from a larger set of designs in a multiple choice recognition format. The Visual Paired Associate subtest was designed as the visual equivalent of the Verbal Paired Associate task in the Wechsler Memory Scale – Revised. The Visual Paired Associate subtest required the recall of the association between colours and six abstract line drawings. The Visual Memory Span task involved tapping a sequence of coloured squares of increasing length in a predetermined order as demonstrated by the examiner. A reversal recall component was also included.

Unfortunately the development of additional visual measures to remedy the predominantly verbal bias of the Wechsler Memory Scale proved unsatisfactory (Lezak et al, 2004). The Figural Memory task was criticised for having greater loading on higher order visual attention span than visual memory and for having no delayed recall procedure (Loring, 1989). As the Figural Memory task assessed recognition memory, comparison in performance to the Visual Reproduction subtest was not possible.

The Visual Paired Associate task clustered more with verbal than non-verbal measures in factor analytic studies and was criticised for the ease with which verbal encoding could be employed (Wong & Gilpin, 1993; Loring et al, 1989). The visual memory span task loaded on memory and attention factors, indicating that processing requirements exceeded the capacity of short-term memory consequently making the measure incomparable to the Digit Span subtest (Bornstein & Chelune, 1988).

A criticism of the Wechsler Memory Scale – Revised was the lack of the provision of cues following failure on the delayed free recall procedure, particularly for the Visual Reproduction measure. Limited recognition procedures were a further shortcoming of the Wechsler Memory Scale – Revised as the relationship between recall and recognition could not be clearly established (Troster et al, 1993; Mayes, 1995). As some clinical groups perform much better on recognition than on recall the lack of cued and recognition procedures limited the ability of the Wechsler Memory Scale – Revised to differentiate among clinical populations (Butters et al, 1988). Researchers have developed cued and recognition procedures for the Visual Reproduction subtest, the Logical Memory subtest and Paired Associate subtest in an attempt to remediate this important omission (Gass, 1995; Fastenau, 1996; Milberg et al, 1986).

1.9.2. Scoring of the Wechsler Memory Scale – Revised

The raw scores of each subtest were weighted and summed to generate an index score utilising age-graded normative tables. The unitary Memory Quotient score from the Wechsler Memory Scale was replaced with five composite index scores; Attention/Concentration; General Memory Index; Verbal Memory Index; Visual Memory Index; and Delayed Recall Index. The index scores were scaled to the same metric system of the Wechsler Adult Intelligence Scale – Revised with means of 100 and a standard deviation of 15 for each age group, allowing comparisons to be made between memory performance and level of intellectual ability (Herman, 1988).

The Attention/Concentration Index comprised the Mental Control, Digit Span and Visual Memory Span subtests. The General Memory Index score is a combination of immediate recall measures of the verbal and visual tasks that also represent separate Verbal Memory and Visual Memory Indexes. The addition of delayed recall procedures for the Logical Memory, Visual Paired Associates, Verbal Paired Associates and Visual Reproduction subtests produced a Delayed Recall Index score (Elwood, 1991). However, the Delayed Recall Index score was not material specific as it encompassed both verbal and nonverbal tasks.

Unfortunately, the General Memory Index cannot be compared to the Delayed Recall Index score as the maximum possible scores that constitute the immediate and delayed indexes differed considerably from one another. The score ranges for the Logical Memory subtest and the Visual Reproduction subtest were greater than the Verbal Paired Associate task and the Visual Paired Associate tasks and there was no delayed recall procedure for the Figural Memory subtest, resulting in differential weightings of the subtests to the index scores. The availability of cues for the Logical Memory subtest in the delayed procedure introduced yet another bias with interpretation of the Delayed Recall Index (Lezak et al, 2004). Furthermore, literal interpretation of what the General Memory Index measured was cautioned against as the index score was derived from heavily weighted measures of verbal memory (Loring et al, 1989; Lezak et al, 2004).

The restriction of major indexes to a low-end score of 50 created floor effects with the potential to inflate memory performance and inaccurately reflect the severity of anterograde amnesia (Butters, et al, 1988). It also limited the discrimination of memory deficits in those with more severe memory disorders (Mitrushina et al, 2005).

Percentile norms were derived for some of the subtests, but no scaled scores were available. This limited the clinical utility of the battery in making useful comparisons between various profiles of clinical populations (Herman, 1988; Naugle et al, 1993).

The inclusion of new subtests and delayed procedures extended the administration time of the battery from approximately 30-45 minutes to an hour. From a practical and clinical viewpoint, this could be a disadvantage with factors such as mental fatigue and reduced concentration potentially confounding performance and subsequent test interpretation.

A short form of the Wechsler Memory Scale – Revised can be administered by eliminating the delayed recall components, but a major criticism of the original Wechsler Memory Scale was the absence of measures evaluating retention of information over time (Wechsler, 1987; Lezak et al, 2004; Mitrushina et al, 2005). A short form comprised of the Logical Memory, Visual Reproduction and Verbal Paired Associate subtests that reduced administration time without sacrificing the clinical efficiency of the measure was described by Woodard and Axelrod (1995) and has received some support (Hoffman, Scott, Tremont, Adams & Oommen, 1997).

1.9.3. Reliability of the Wechsler Memory Scale – Revised

Reliability coefficients for the subtests and indices across age groups ranged from .41 to .90 with a median of .74. Ivison (1990) reported that the internal consistency for Wechsler Memory Scale – Revised subtests ranged from .45 to .79 and for the General Memory Index it was .80, comparable to that documented in the manual. Restricted score ranges however, resulted in low reliability for several of the subtests. Inter-scorer reliability for the Logical Memory subtest and the Visual Reproduction subtest was .99 and .97 respectively (Wechsler, 1987).

The lack of alternative forms available for times of serial administration was also a significant limitation of the Wechsler Memory Scale – Revised. However, independent tables regarding reliable change have been developed for consultation (Lezak et al, 2004).

1.9.4. Normative Data of the Wechsler Memory Scale – Revised

The normative data was extended to include information from 316 individuals aged 16 years, 0 months to 74 years, 11 months. Approximately 50 subjects were included in each age group stratified at six levels: 16-17, 20-24, 35-44, 55-64, 65-69 and 70-74. The normative age groups were set to parallel those of the Wechsler Adult Intelligence Scale –Revised age intervals for comparative purposes. Although the manual reported norms stratified at nine age levels, the norms for three of the groups 18-19; 25-34, and 45-54 were estimated statistically by interpolation.

The normative sample was stratified according to age, gender, ethnicity, geographic region, IQ and education (at three levels 0-11 years, 12 years and 13 years or more). However, no normative data were provided by years of education and IQ level, despite this information being readily available (Mitrushina et al, 1999).

Despite the significant improvement in the standardization and normative data, large standard errors of measurement along with sampling errors were found on the Wechsler Memory Scale – Revised. Major criticisms were made regarding the use of interpolated means of the normative data that were based on a relatively small standardization sample (Mitrushina et al, 2005; Loring, 1989). The inclusion of actual performance data for the 55-64 age band would have been advantageous as memory abilities of individuals in this age group can vary considerably and it is often the age when memory difficulties are first identified (Butters et al, 1995; Lezak et al, 2004).

Furthermore, the evaluation of memory problems that arise later on in life due to neuro-degenerative processes such as dementia were limited with normative data only available to the age of 74. Ivnik, Malec, Smith, Tangalos, Petersen, Korkmen and Kurland (1992) remedied this problem by publishing norms for individuals aged 56 to 94 years as part of Mayo’s Older American Normative Studies (MOANS). Unfortunately, the MOANS indices differ from the Wechsler Memory Scale – Revised so the summary scores are not interchangeable.

Additional normative studies have been conducted to account for some of the standardization inadequacies. These studies have largely been based on American populations (Mitrushina et al, 2005). However, Shores and Carstairs (2000) published normative data for an Australian population aged 18 to 34 years providing local norms for this age group.

1.9.5. Validity of the Wechsler Memory Scale – Revised

Wechsler reported general memory and attention factor solutions to best represent the structure of the Wechsler Memory Scale – Revised in a mixed clinical group. A subsequent confirmatory factor analysis reportedly favoured a similar two-factor solution (Roid et al, 1988). Bornstein and Chelune (1988) replicated Wechsler's analyses utilising immediate measures and identified similar general memory and attention factors. The inclusion of both immediate and delayed indexes revealed a three-factor structure of discrete attentional, verbal and visual factors (Chelune & Bornstein, 1988).

Roth, Conboy, Reeder and Boll (1990) also derived a three-factor solution of attention, general memory and delayed recall factors in a traumatic brain injured population. Burton, Mittenberg and Burton (1993) confirmed the factor solutions described by Roth et al (1990) as the most accurate in explaining the variability in memory function. Support for the three-factor model has been demonstrated in clinical studies including Alzheimer's disease, Huntington's disease, Wernicke Korsakoff Syndrome and severe closed head injury populations (Butters et al, 1988; Wechsler 1987). Other authors have found the factor structure of the Wechsler Memory Scale – Revised varied according to age, years of education and the diagnostic group (Loring et al, 1989; Bornstein & Chelune, 1988).

More recently, joint analysis studies of the Wechsler Memory Scale – Revised together with the Wechsler Adult Intelligence Scale – Revised have supported five and six factor models that included verbal and visual memory factors (Leonberger, Nicks, Larrabee &

Goldfader, 1992; Smith, Ivnik, Malec, Petersen, Tangalos & Kurland, 1992; Bowden, Carstairs & Shores, 1999; Larrabee, 2000)."

1.9.6. Clinical Utility of the Wechsler Memory Scale – Revised

Despite weaknesses, the Wechsler Memory Scale – Revised did represent a clear improvement over the previous version due mostly to the inclusion of a Delayed Recall Index and the separation of attention measures from the General Memory Index. The Wechsler Memory Scale – Revised was found to be a valid measure in the assessment of severe anterograde memory disorders such as those with amnesic, cortical and subcortical dementias (Butters et al, 1988).

The development of separate verbal and visual indexes was anticipated to demonstrate sensitivity to lesion laterality, but this has not been consistently reported. In a group of patients with temporal lobe resection, summary indices could not distinguish lateralized verbal and non-verbal deficits (Loring et al, 1989). Furthermore, in a study of 60 unilateral temporal lobe seizure patients, Naugle, Chelune, Cheek, Luders and Awad (1993) found a significant drop on measures of verbal memory in patients with left temporal lobotomies, but no decrements in nonverbal memory were found in those with right temporal lobotomies both prior to and following temporal resection. In fact, there was no indication right temporal lobotomy resulted in reduced performance on immediate or delayed procedures relative to preoperative baseline or to non-surgical patients with intractable epilepsy. However, patterns of non-verbal memory impairments have been reported to emerge as a within subject effect in a study of a variety of unilateral brain damaged patients (Chelune & Bornstein, 1988).

1.10. The Wechsler Memory Scale – Third Edition

In its second revision, the Wechsler Memory Scale underwent major changes to reflect theoretical and experimental advances in the assessment of memory. Changes to the structure, content, scoring, and administration of the test resulted in a very different measure to its predecessors (Tulsky, Chiaravalloti, Palmer & Chelune, 2003).

1.10.1. Structure and Content of the Wechsler Memory Scale – Third Edition

The second revision of the Wechsler Memory Scale has six primary subtests: (1) Logical Memory; (2) Family Pictures; (3) Verbal Paired Associates; (4) Faces; (5) Visual Memory Span (now called Spatial Span); (6) Letter-Number Sequencing; and five optional subtests: (1) Information and Orientation; (2) Mental Control; (3) Digit Span; (4) Visual Reproduction; and (5) Word Lists.

Three of the six primary subtests (i.e., Logical Memory, Verbal Paired Associate and Spatial Span) were retained from the Wechsler Memory Scale – Revised, but were modified significantly. Minor changes to the wording of the first story of the Logical Memory subtest occurred, but an entirely new second story was developed. Due to the potential bias of the second story to evoke an emotional response in patients who had a motor-vehicle accident or experienced trauma it was eliminated (Tulsky & Ledbetter, 2000). An additional administration procedure of the second story was included in an effort to assess the learning of prose from repeated exposure to the material. A two-choice recognition procedure was also devised for administration following the delayed recall procedure for both stories (Lezak et al, 2004). A revision of the scoring system of the Logical Memory included the scoring of thematic content (major themes of the story) in addition to rewarding points for accuracy of story recall.

The Verbal Paired Associate subtest was entirely revised with the elimination of the easy word pairs and administration of eight new low association word pairs, plus the addition of four trials and a recognition procedure. A major clinical disadvantage of the revised Verbal Paired Associate measure was the elimination of the easy word pairs that provided patients an opportunity to succeed and used to alert the examiner of possible poor task motivation if failure occurred on those items. A low ceiling effect was a criticism of the recognition procedure, as correct word pairs always appeared together and only one word from the pair was required to be recognised for a response to be awarded (Lezak et al, 2004; Tulsky & Ledbetter, 2000).

The Visual Memory Span card of a two dimensional configuration was replaced with a three dimensional board containing numbered cubes for increased ease of administration and scoring. The task involves tapping a sequence of visual spatial locations in the same order as the examiner, and in a reverse sequence order.

Three newly developed measures were included to make up the six core subtests of the Wechsler Memory Scale – Third Edition, namely Letter-Number Sequencing, Faces, and Family Pictures subtests. The Letter-Number Sequencing task involved the presentation of a letter and number series of increasing length (from two to eight elements) required to be re-organised in an ascending and alphabetical sequence.

The Visual Paired Associates and Figural Memory subtests of the Wechlser Memory Scale – Revised were eliminated from the Wechsler Memory Scale – Third Edition as clinical and research data did not support the validity of these subtests as measures of non-verbal memory (Lezak et al, 2004). These measures were replaced with the Faces subtest and Family Pictures subtest as measures of visual memory.

The Faces subtest was developed as a visual recognition memory task, with both immediate and delayed components. It utilised 24 target faces, shown one at a time for approximately two seconds each. The target faces were recognised from a total set of 48 faces interspersed amongst an equal number of distracters.

The Faces measure was included because of evidence that facial recognition tasks have shown differential specificity to right versus left hemisphere damage (Tulsky, 2004; Tulsky et al, 2003). Furthermore, faces represent a unique type of visual stimulus less amenable to verbal encodes. Given the measure represents a task in everyday life for individuals, it was thought to have ecological validity, addressing a criticism of the Wechsler Memory Scale – Revised visual memory subtests.

The Family Pictures subtest was developed specifically for the Wechsler Memory Scale – Third Edition and was new to clinical practice and research fields alike (Wechsler, 1997). The Family Pictures subtest was developed as an analogue to Logical Memory subtest, as it was a measure of complex and meaningful visual information. The measure involved the presentation of four scenes including members of a family engaged in a number of activities, shown for 10 seconds in a sequential order. Hence, recall was only examined after all four pictures were presented. Memory was evaluated using structured questions regarding which family members were involved in each scene, the type of activity that was undertaken and their spatial location on a 2x2 grid. Delayed recall was evaluated after a 25 to 35 minute interval.

The combination of the Faces subtest and Family Pictures subtest to make the Visual Memory Index in the Wechsler Memory Scale – Third Edition has been extensively criticised. The Faces subtest did not have a strong correlation with the Family Pictures measure, suggesting the two subtests evaluate different aspects of memory (Lichtenberger, Kaufman & Lai, 2002). Little communality shared between these measures presented problems in the interpretation and raised concerns about the construct validity of the Visual Memory Index score (Millis, Malina, Bowers & Ricker, 1999; Tulsky, Chelune & Price, 2004). These concerns led to the suggestion of substituting the Faces subtest with the Visual Reproduction measure in the generation of a Visual Memory Index. Analysis of the combination of these measures revealed higher correlation coefficients with other indexes and comparable clinical sensitivity to the original scores, resulting in a better measure of visual memory (Hawkins & Tulsky, 2004).

Also the Family Pictures subtest had a stronger correlation with Logical Memory than the Faces subtest. This suggested the Family Pictures subtest was highly available to verbal encoding processes and was not surprising considering the mode of investigation of recall was predominantly verbal. The presentation of all four pictures before evaluating immediate free recall has been criticised as assessing not only immediate memory, but also post-interference recall and delayed recall (Lezak, 2004).

The core Visual Reproduction and Digit Span measures of the Wechsler Memory Scale – Revised were changed to optional subtests in the latest revision without a rationale provided in the manual (Horton, 1999). The Word List subtest was a verbal learning task also included as an optional measure in the Wechsler Memory Scale – Third Edition. The Word List subtest was based on the Rey Auditory Verbal Learning Test and Hopkins Verbal Learning Tests and it involved the presentation of a list of 12 semantically unrelated words repeated over four trials, with recall assessed after each trial. A second list was then introduced to act as an interference task. Recall of the first list was re-appraised and delayed recall evaluated after a 25 – 35 min interval. A recognition procedure followed delayed recall where the initial 12 words were interspersed amongst 12 distracters.

1.10.2. Scoring of the Wechsler Memory Scale – Third Edition

A fundamental shift in the processing of raw scores was made in the Wechsler Memory Scale – Third Edition. Almost all raw scores can be converted to age adjusted scaled scores (with a mean of 10 and a standard deviation of 3) from the use of specific tables unique for each age group (Lezak et al, 2004). This allows for an equal contribution of each subtest to account for the index scores and for reliable comparisons between indices to be available.

Age-adjusted scaled scores were developed for the optional subtests, making them comparable to the index scores. The conversion of scaled scores into percentiles was also available for most subtests (Wechsler, 1997). As the raw score distributions of the Information and Orientation subtest and the Visual Reproduction discrimination total score were highly skewed, the use of percentiles over scaled scores were favoured for these measures (Lezak et al, 2004).

The core tests generated eight primary indices: Auditory Immediate, Visual Immediate, Immediate Memory, Auditory Delayed, Visual Delayed, Auditory Recognition Delayed, General Memory, and Working Memory. Each index has a mean of 100 and a standard deviation of 15, making it comparable to the Wechsler Adult Intelligence Scale – Third Edition.

The Wechsler Memory Scale – Third Edition provided the computation of four auditory process composites derived from immediate and delayed performances on the Logical Memory and Verbal Paired Associate subtests. The Single Learning trial score was for recall after an initial presentation of information, the Learning Slope measures improvements in performance over trials, Retention was a measure of the ability to retain material over the delay interval, and Retrieval documents the differences between free recall and recognition scores. Confidence intervals for test-retest measurement errors were provided for guidance in interpreting change between test administrations (Lezak et al, 2004). Discrepancy analyses between index scores, scaled scores and Wechsler Adult Intelligence Scale – Third Edition subtest scores were also derived (Wechsler, 1997).

The Wechsler Memory Scale – Third Edition saw a fundamental shift in the conceptualization of memory. The General Memory Index was based on delayed recall measures and consistent with clinical approaches, that performance after a time delay best represents memory function. The General Memory Index of the Wechsler Memory Scale – Revised and the Memory Quotient of the Wechsler Memory Scale were based purely on immediate recall measures, consequently making the measures incomparable. However, the General Memory Index of the Wechsler Memory Scale – Third Edition was a combination of delayed recall and recognition performances. This has been heavily criticised given substantial evidence that free recall can be compromised and recognition can be relatively intact in many memory disorders (Hawkins, 1998).

The availability of both delayed recall and recognition indices was an important inclusion in the Wechsler Memory Scale – Third Edition in distinguishing deficits in retrieval from inefficient learning and consolidation. However, the Auditory Recognition Index has been criticised for low ceiling effect, limiting the utility of the index in the detection of deficits in those other than the severely memory impaired (Lezak et al, 2004).

Furthermore, problems in scoring the individual subtests of the Wechsler Memory Scale –Third Edition were identified. For example, no rationale was provided for the differential weighting of the two stories on the Logical Memory subtest. On the recognition procedures, the score on the Logical Memory subtest was 50% chance. On the Verbal Paired Associate subtest, both words in the pair were provided and some word pairs were presented twice, with no justification. The recognition measure of the Visual Reproduction subtest incorporated elements of different designs, which lead to confusion as to which design was being evaluated. Despite problems in the structure of the Wechsler Memory Scale – Third Edition that impact on the clinical utility of the measure a number of features represented a clear improvement on the previous editions.

1.10.3. Reliability of the Wechsler Memory Scale – Third Edition

Test-retest reliability for the core subtests ranged from .62 to .82 with median of .74 for a retest interval of 35.6 days. Split-half coefficients for subtests across all age groups ranged from .74 (Faces I and II) to .93 (Verbal Paired associates I) (Wechsler, 1997).

1.10.4. Normative Data of the Wechsler Memory Scale – Third Edition

The normative data for the Wechsler Memory Scale – Third Edition was a significant improvement to the norms available from the Wechsler Memory Scale – Revised. The Wechsler Memory Scale – Third Edition norms were collected from 1250 individuals aged 16 – 89 years that constituted half the total sample collected during the standardisation of the Wechsler Adult Intelligence Scale – Third Edition (Wechsler, 1997). The Wechsler Memory Scale – Third Edition and the Wechsler Adult Intelligence Scale – Third Edition were co-normed to provide clinically useful information about a wider range of cognitive functions allowing for more meaningful and comparative interpretations (Tulsky et al, 2003).

The standardisation battery used to norm the Wechsler Memory Scale – Third Edition was not equivalent with the published version. Specifically, the standardisation battery included the administration of additional measures between the immediate and delayed recall procedures of the core subtests that were not included in the published Wechsler Memory Scale – Third Edition battery for psychometric or conceptual reasons. The shortening of the measure raised concerns about variable levels of fatigue on performance, the reordering of the test sequence, and the interference effects potentially created by administration of the standardisation versus published battery (Doss, Chelune & Naugle, 2000). However, the validity of the normative data has been supported (Doss et al, 2000; Zhu & Tulsky, 2000; Holdnack, Lissner, Bowden & McCarthy, 2004).

1.10.5. Validity of the Wechsler Memory Scale – Third Edition

The Wechsler Memory Scale – Third Edition Technical Manual (Wechsler, 1997) reported factor analyses supporting three and five factor solutions with separate factors for immediate and delayed recall. A three factor model best fit the data for the standardization sample aged between 16 and 29 years and a five factor solution best fit the 30 to 64 and 65 to 89 age groups suggesting the factor structure might change as a function of age.

Millis, Malina, Bowers and Ricker (1999) attempted to replicate the factor structure that was reported in the Wechsler Adult Intelligence Scale – Third Edition and Wechsler Memory Scale – Third Edition Technical Manual, but were unable to achieve convergence on the five factor model. A three-factor model of working memory, verbal memory and visual memory was documented to best fit the standardization sample.

Similarly, Price, Tulsky, Millis and Weiss (2002) supported a three-factor model of immediate and delayed verbal memory, immediate and delayed visual memory, and working memory to represent the factor structure of the Wechsler Memory Scale – Third Edition. In a temporal lobe epilepsy population, a general memory factor that included both verbal and visual memory was the best fit (Wilde et al, 2003). Other factor analytic studies of the Wechsler Memory Scale – Third Edition also supported verbal and visual memory factors (Tulsky & Price, 2003; Larrabee, 1999; Gladsjo, McAdams, Palmer, Moore, Jeste & Heaton, 2004; Tulsky, Ivnik, Price & Wilkins, 2003).

Currently, separate immediate and delayed memory factors are not supported, hampered by specification errors (Tulsky, et al, 2003). Further research on the factor structure of the Wechsler Memory Scale – Third Edition utilizing clinical populations may reveal separate immediate and delayed constructs, as memory performance may vary as a function of cerebral compromise (Price et al, 2002).

1.10.6. Clinical Utility of the Wechsler Memory Scale – Third Edition

Direct comparisons between the sensitivity of the Wechsler Memory Scale - Revised (WMS-R) and Wechsler Memory Scale – Third Edition (WMS-III) are difficult. This is in part due to the absence of IQ data in the WMS-R manual and lack of information regarding the severity of the disorders for the clinical groups included. At the time of release, published data in the manual of the WMS-III on the performance of patients with neurological disorders was limited. However, preliminary evidence of the WMS-III suggested that it is at least as sensitive to neurological dysfunction as the WMS-R in the assessment of Alzheimer’s disease patients (Horton, 1999; Tulsky et al, 2003). Further studies on these client groups utilizing the WMS-III are necessary to establish the utility of the measure in a variety of clinical populations (Millis et al, 1999; Tulsky 2004).

The Wechsler Memory Scale – Third Edition manual reported an administration time of 30 to 40 minutes for the entire battery. However, much longer administration times of up to two hours have been reported in population groups with known memory impairments (Axelrod, 2001; Lichtenberger et al, 2002).

1.11. Memory Batteries versus Individual Tests

A potential problem with utilising a memory battery can be the assumption that the entire battery will be administered at a given time. In a practical sense, not all tasks in a memory battery may be required as particular subtests may not be relevant to the patient’s reported difficulties. Moreover, the administration of an entire battery may preclude the assessment of other important cognitive functions specific to the referral question. Furthermore, the time required for the administration of an entire memory battery can be lengthy, which is often not viable due to the demand for quick and efficient assessments in most clinical settings.

Fatigue often poses a significant risk to mental endurance in lengthy assessments in many clinical populations and the elderly. As such, the administration of individual tests in a flexible battery approach provides a more clinically useful means of assessment in addressing the referral question in certain cases (Lacritz & Cullum, 2003). The use of tests from different sources does not enjoy the advantage of identical normative data as that from a fixed test battery, but psychometric information (such as standard deviations, percentiles, raw scores accompanied by their statistical descriptions) can allow for some comparisons to be made, although interpretation should be approached cautiously.

Clinically useful measures of verbal memory have been developed via the Wechsler Memory Scale and other tests. The development of non-verbal measures has progressed more slowly. Arguably one of the most well-known and well used measures of non-verbal memory has been the Visual Reproduction subtest from the Wechsler Memory Scale and its revisions.

1.12. The Visual Reproduction subtest of the Wechsler Memory Scales: A Closer Look

1.12.1. Design of the Visual Reproduction subtest of the Wechsler Memory Scale and its Revisions

The Visual Reproduction subtest of the Wechsler Memory Scale was originally developed as a measure of immediate recall. Two variations of the Visual Reproduction subtest were developed, but due to the lack of psychometric data available for Form II, only Form I was used clinically.

Studies on the factor loading of the Visual Reproduction subtest have shown it has a strong relationship with the Visual Organization factor of the Wechsler Adult Intelligence Scale, calling into question the Visual Reproduction subtest as a measure of visual memory (Lezak, 1995; Larrabee & Curtiss, 1995; Williams, Rich, Reed, Jackson, LaMarche & Boll, 1998). In addition, no information was provided on influence of education on either immediate or delayed recall. The less educated individuals were found to perform more poorly on the measure than those who had a higher level of education (Ardila & Rosselli, 1989). Practice effects were also identified with a group of patients aged 69, having gained almost two points on retesting a year later, with gains disappearing upon retesting the following year.

The Visual Reproduction subtest of the Wechsler Memory Scale was criticised for the lack of delayed recall procedures. The inclusion of a delayed recall procedure in the Visual Reproduction subtest of the Wechsler Memory Scale – Revised made it a more useful measure of visual memory than the original Wechsler Memory Scale as delayed recall provides a better measure of memory function (Larrabee & Curtiss, 1995; Lezak et al, 2004).

The Visual Reproduction subtest has demonstrated a strong loading on both verbal and visual memory factors. Bornstein and Chelune (1989) found the Visual Reproduction subtest had a stronger relationship with verbal memory in older (40-55 and 56+) than the younger age groups (below 40 years). The findings suggest that older individuals may rely more on verbal strategies in the recall of non-verbal material than younger individuals (below 40 years). This is potentially a reflection of reduced fluid intellectual abilities in older adults and the memory changes that are associated with aging (Bornstein & Chelune, 1989; Fahle & Daum, 1997).

Due to the difficulty in establishing a measure of non-verbal memory without confounds of other cognitive processes such as visuo-perceptual, constructional and verbal elements, studies of non-verbal memory are often flawed (Heilbronner, 1992). In addition, the use of simple and familiar geometric figures has been criticised as easily available to verbal encoding processes, thus further confounding performance and potentially resulting in non-verbal memory impairments going undetected (Eadie & Shum, 1995).

One way of minimising the interaction effect between verbal and non-verbal processes is to develop a test with unfamiliar and complex geometric shapes that would be difficult to encode verbally (Eadie & Shum, 1995). Greater complexity and exposure to novel designs is linked to right hemispheric lateralization and non-verbal processes (Redoblado, Grayson & Miller, 2003; Jones-Gotman, Zatorre, Olivier, Andermann, Cendez, Stuanton, et al, 1997). However, there is a possibility that complex designs will make them too difficult for both brain damaged and non-brain injured individuals to recall. Furthermore, the most complex can be verbalised to some extent.

The complete elimination of verbal processes is not assumed possible or even desirable, as the use of verbal strategies or the application of semantic processes to assist with recall provides clinically useful information in the way material may be organised and remembered (Heilbronner, 1992). Nonetheless, the development of a measure with minimal confounds of other cognitive processes to determine the relative contribution of non-verbal memory function to the overall nature of memory impairment is important.

The Visual Reproduction subtest of the Wechsler Memory Scale – Revised was criticised for the lack of recognition and copy trials to identify any motor control difficulty as part of the standard administration of the measure (Gfeller, et al, 1995; Lezak et al, 2004). As the Visual Reproduction subtest requires a visuo-motor response (i.e., drawing), the assessment of visual-spatial, constructional and motor control problems is considered clinically important to determine its contribution and to rule out any component processes that can significantly influence performance (Haut, Weber, Demarest, Keefover & Rankin, 1996; Moye, 1997).

The Boston Revision adequately addressed the lack of a recognition procedure and matching trials of the Wechsler Memory Scale – Revised, but problems with its administration were identified. The administration of the recognition procedure immediately following an initial presentation allowed the potential for learning to confound performance on delayed recall and consequently could not be supplemented as part of the memory battery (Mitrushina et al, 2005). Russell's revision (1975) was also criticised as a point was subtracted from the total score if a prompting procedure was utilised with the potential to confound recall (Loring & Papanicolaou, 1987). Fastenau (1996) developed recognition, perceptual matching and copy trials for the Visual Reproduction subtest. However, poor psychometrics, limited reliable data, and the presence of ceiling effects due to small range of test items did not enhance the utility of the Visual Reproduction subtest and its clinical application.

Due to the criticisms of the Wechsler Memory Scale – Revision version, the Wechsler Memory Scale – Third Edition incorporated recognition, perceptual matching and copy procedures (Tulsky et al, 2003). However, the content and format of the recognition task has raised potential problems. Recognition memory is evaluated using a series of 48 designs with 14 target designs and 34 distracters. No rationale was provided regarding the distribution of target versus distracter designs. As recognition memory is assessed via a yes/no format, this provides a 50-50 probability of a correct response raising the question of guessing scores to confound true performance.

In addition, parts of the target designs are presented as separate items and most of the distracter items are similar to the target designs, leading to confusion and increasing the likelihood of incorrect responding due to partial recognition (Dowling, 1998). Furthermore, a 'no' response to all recognition items will still yield a score that falls within the normal range for individuals aged 80-89. From a practical viewpoint, the inclusion of 48 items makes the task lengthy and somewhat time consuming, which can be problematic for clinical populations where fatigue and reduced concentration are common issues.

The second revision of the Visual Reproduction subtest was also criticised for the lack of a cued recall procedure as an intermediate step between the assessment of delayed free recall and recognition. If a partial cue is sufficient to trigger recall, this can provide useful information on the severity of a retrieval deficit. Although test-retest computations are available for the Wechsler Memory Scale – Third Edition, no alternative forms for use in serial assessments were developed (Lineweaver & Chelune, 2003).

Potential problems with the development of the Visual Reproduction designs of the Wechsler Memory Scale – Third Edition have been identified. Design A was criticised as it is actually half of Design B (Lacritz & Cullum, 2003). In addition, three of the designs have overlapping features (e.g. include flags as elements), consequently confounding performance and potentially making it difficult to distinguish perseveration. This is further problematic as it may be difficult to differentiate what item is being recalled on delayed recall, and subsequently lead to confusion in the scoring and interpretation of delayed recall performance.

1.12.2. Scoring the Visual Reproduction Designs of the Wechsler Memory Scale and its Revisions

The lack of explicit scoring procedures for the original Visual Reproduction designs allowed for large score discrepancies to occur, mostly arising from differences of opinion about the degree of accuracy required (Lezak et al, 2004).

The rules for the revised scoring system of the Visual Reproduction subtest of the Wechsler Memory Scale – Revised were expanded and made more explicit. The addition of detailed drawing examples provided further clarity of the scoring principles. Despite improvements, several criticisms and problems with the scoring system remained. Although a high level of inter-scorer agreement on the Visual Reproduction subtest of the Wechsler Memory Scale – Revised was reported to be a high .97 (Wechsler, 1987; McGuire & Batchelor, 1998), the potential for significant disagreement between scorers has been reported.

Wechsler (1987) reported a scoring difference of a maximum of four points on immediate and delayed recall scores on the Visual Reproduction subtest with an average difference of 1.50 points. Even though the discrepancy did not exceed four points for each design and such a difference was infrequent, this equated to a possible 16-point discrepancy on the total score, suggesting a need for further refinement of the scoring system. Furthermore, variability in internal consistency estimates and reliability of the scoring criteria ranging from .46 to .71 on immediate recall and .38 to .59 on delayed recall trials have been documented (Williams et al, 1998).

Furthermore, for participants aged 20 – 64 the stability coefficients ranged from .56 to .80, with an average of .68 on the immediate recall, and a range of .58 to .68 with an average of .63 on delayed recall for a 4-6 week test-retest period (Wechsler, 1987). Similar stability coefficients of .80 or greater have been reported at 4-5 month retest interval with an average gain of 2 – 6 points on the Indexes of the Wechsler Memory Scale – Revised (Bowden, Whelan, Long & Clifford, 1995).

The scoring range of the Visual Reproduction subtest of the Wechsler Memory Scale – Revised was increased from a maximum score of 14 to 41 (Wechsler, 1987; Smith, Malec & Ivnik, 1992). Despite the extended score range of the subtest, each design had differential weighting and an unequal contribution to the total score. Even though the designs potentially contained differing number of elements as a reflection of the complexity of the designs, no rationale for the differential weighting was provided in the manual (Wechsler, 1987).

The limited range of scores for the first three designs collectively contributed to approximately 50% of the total score (i.e. Cards A and B have a maximum seven points, Card C has a maximum of nine and Card D has a maximum of 18 points). This placed undue importance on performance on the fourth design. A poor score on Card D could result in a total score below the 25th percentile potentially underestimate memory ability that can have significant clinical implications.

If all designs were equally weighted, the total score could potentially be prorated when performance one design was disrupted or invalidated, by utilising the scores on the other designs. Moreover, equal weighting of designs could allow comparisons across designs to be made and potentially provide important clinical information between different clinical groups.

The restricted range of scores on the Visual Reproduction subtest of the Wechsler Memory Scale – Revised can lead to less discrimination between individuals and contribute to more ceiling and floor effects, increases the error score and reduce the discriminatory power of the measure. Specifically, the scoring criteria has been criticised for floor and ceiling effects prominent in the elderly and mildly impaired clinical populations respectively (Lezak et al, 2004).

While the scoring criteria for the Visual Reproduction subtest of the Wechsler Memory Scale – Revised were more explicit than the original version, the precise measurement of angles, distances and ratios potentially compromises the construct validity of the subtest by overemphasizing visual-perceptual motor skills and spatial reasoning abilities (Williams et al, 1998). In addition, the administration procedures were expanded but the instructions do not include a caution to draw carefully, as carelessness and impulsivity would have a negative impact on the score. For example, Card C scoring guide item 4 states ‘No dot may be a circle,’ drawing one dot as a circle would result in no credit awarded even though all elements may be present (Lacritz & Cullum, 2003). Consequently, poor attention to detail may result in a deficient performance and reflect a memory impairment that does not actually exist (McGuire & Batchelor, 1998). A good scoring system from a clinical perspective should provide tolerances for carelessness, impulsivity or poor motor control, and award points for memory recall rather than artistic ability and precision.

Furthermore, the lack of credit for the recall of partial features of the designs has the potential to underestimate memory performance. The inclusion of criteria for awarding partial recall in addition to items that credit exact reproductions of designs can provide a greater range in the recall quality and increase the discriminatory ability of the subtest (Clark, 2000).

A further criticism of the Wechsler Memory Scale – Revised scoring system is the lack of scoring each item independently. For example, on Design C if failure on item five occurs then items six to nine are also scored zero. The dependence of one item on another also implies an assumption to how information in memory is stored. There is certainly no clinical information to support this and there was no rationale for linking criterion reported in the manual. The development of a system that awards credit for each item independently would eliminate any assumptions regarding how information was processed, providing a better estimate of memory performance.

The scoring criteria has also been scrutinised as primarily verbal and based on semantic prompts, thus providing greater opportunities for verbal encoding to mediate non-verbal memory processes (Smith et al, 1992). Interestingly, researchers have implicated laterality differences to the duration of exposure to material, with longer periods allowing greater use of semantic encoding and verbal strategies (Iidaka et al, 2000). Developing a scoring system that focuses more on the relationships between shapes and elements of designs may better reflect non-verbal aspects of visual stimuli. Indeed the scoring protocols developed to evaluate the quality and nature of memory recall may influence its sensitivity and specificity as a measure of non-verbal memory.

The Wechsler Memory Scale – Third Edition expanded the scoring system with detailed criteria and the availability of partial credit for most items, extending the range of the test to a maximum of 104 points (Wechsler, 1997). The inclusion of new and altered designs improved the issue of floor and ceiling effects a criticism of previous editions. The availability of subtest scaled score conversions for the nine age groups in the standardization sample allowed for performances across age groups to be compared

(Lezak et al, 2004). Although substantial improvements in the psychometrics of the Wechsler Memory Scale – Third Edition were evident, problems associated with the designs, recognition procedure and scoring system leave significant doubts about it being a useful measure of non-verbal memory (Dowling, 1998).

1.13. Rationale of the Current Study

The wealth of research and clinical data available on the Visual Reproduction subtest of Wechsler Memory Scale – Revised still makes it a popular measure amongst many clinicians in the neuropsychological assessment of memory. In practical terms, the Visual Reproduction subtest of the Wechsler Memory Scale – Revised takes a very short time to administer in comparison to the third edition. This is important given the demand for quick and efficient neuropsychological evaluations in most clinical settings. Furthermore, recent normative data on an Australian population of 399 adults aged 18-34 years the Wechsler Memory Scale – Revised holds a definite advantage to its use in Australia, with more culturally appropriate and demographically corrected norms (Waldman & Dickson, 1991; Shores & Carstairs, 2000; Carstairs & Shores, 2000).

In an attempt to improve the Visual Reproduction subtest of the Wechsler Memory Scale – Revised as a measure of non-verbal memory, Clark (2000) modified the scoring system and administration procedures. The Revised Scoring System was developed on the following guidelines and principles briefly outlined below:

- Each of the four designs would generate a score out of 20 points.
- Each figure will be given equal weighting and assigned a similar number of criteria.
- The criteria would address as many different aspects of the design rather than a few key elements.
- Criteria would be scored independently from one another
- Particular attention would be given to spatial aspects and relationships between elements of the designs

- Imperfect recall of the designs would still score points.
- Criteria would be allocated for perfect reproductions of the designs.
- Criteria would be allocated for perfect reproductions that did not contain any extra elements.
- Tolerances for angles, line lengths, minor gaps and curves would be included.
- A grading of difficulty for items would be developed in an attempt to assist with the marking process.
- The number of criteria that addressed the most frequently reproduced aspects of the design would be limited.

For detailed description of the guidelines and principles see Design section 2.3.3.

The Revised Scoring System developed by Clark (2000) demonstrated reliability at least equal to that of the Original Scoring System. A high correlation between the scoring systems suggested a similar grading of memory performance, but the Revised Scoring System had the advantage of a greater range of scores. Clark (2000) also developed a Non-verbal Index to provide an indicator of deficient non-verbal memory, which preliminarily discriminated between persons with left and right hemisphere lateralised lesions, providing support as a diagnostic instrument. Normative data is currently being collected (H. Madill, personal communications, July 18, 2007).

Overall, the preliminary evidence of the Revised Scoring System developed by Clark (2000) appeared to address a number of the limitations of the Original Scoring System of the Visual Reproduction subtest of the Wechsler Memory Scale – Revised. However, a criticism not addressed by previous research was the lack of available alternative forms for the Wechsler Memory Scale – Revised subtests for times when serial assessments are required.

Serial assessments are important and conducted frequently to evaluate rate of recovery from injury in acute periods, determining the effectiveness of rehabilitation and therapeutic interventions post-acute and in long-term follow-up. The problem of practice effects on repeat administration of neuropsychological measures is particularly salient in the domain of memory (Lineweaver & Chelune, 2003). The advantage of having alternative forms available for widely used memory tests is to minimise the potential for practice effects and learning of material that might confound performance and subsequent interpretation of test scores (Benedict & Zgaljardic, 1998).

Although the potential for individuals to develop strategies to perform better on a measure is probably unavoidable in serial assessments as a result of familiarity with the testing procedure, the potential for practice-related measurement error is increased (Lineweaver & Chelune, 2003). Moreover, the measurement error as a result of practice effects from repeated examination using the same measure is even more problematic. On repeat examination, the previous exposure to the stimuli is likely to inflate the immediate recall score and confound test performance (Lineweaver & Chelune, 2003; Moye, 1997).

Indeed, the study by Theisen, Rapport, Axelrod and Brines (1997) highlighted a significant increase in performance over a two week test-retest interval on the Visual Reproduction subtest of the Wechsler Memory Scale – Revised amounting to two scaled score points. Moreover, the increase in performances on non-verbal measures has been identified to be greater over longer test-retest intervals than the practice effects found using verbal memory measures (Benedict & Zgaljardic, 1998). Therefore, score gains due to practice effects can have significant diagnostic implications regarding memory function and result in misleading interpretations about improvement in function or the benefit of therapeutic interventions. This increases the importance of having alternative forms available for non-verbal memory measures for clinical use.

1.13.1. Aims and Hypotheses

The first aim was to develop an alternative form of the Visual Reproduction subtest of the Wechsler Memory Scale – Revised. This would include the addition of cued and recognition procedures to supplement the alternative form, in order to increase the potential clinical utility of the measure.

The second aim of the study was to develop a scoring system for the alternative form of the Visual Reproduction subtest of the Wechsler Memory Scale – Revised. For the purposes of this study, it is called the Alternative Scoring System and the Alternative Form.

The Alternative Scoring System would be devised using the principles adopted in the Revised Scoring System developed by Clark (2000) in the revision of the Original Scoring System of the Visual Reproduction subtest of the Wechsler Memory Scale – Revised.

The hypotheses of this study are as follows:

1) The Alternative Form would have a moderate-high positive correlation with the Visual Reproduction subtest of the Wechsler Memory Scale –Revised.
2) The Alternative Form would demonstrate moderate-high internal reliability.
3) The Alternative Form would demonstrate convergent validity with moderate-high positive correlation with the Rey Complex Figure Test, a measure of non-verbal memory.
4) The Alternative Form would demonstrate discriminant validity with weak correlations with the Logical Memory subtest and the Rey Auditory Verbal Learning Test, measures of verbal memory and learning.
5) The Alternative Form would demonstrate construct validity with weak – moderate correlations with general intellectual ability using the Wechsler Adult Intelligence Scale – Revised.
6) The development of cuing and recognition procedures for the Alternative Form would provide additional information regarding memory performance.

METHOD

2.1. Participants

The participants who volunteered to take part in this study were recruited from metropolitan Melbourne and surrounding areas via convenience sampling. The participants were 44 non-clinical adults aged between 25 and 51 years with a mean age of 37.93 (SD = 7.60). Twenty-four participants were male (55%) and 20 were female (45%), with 91% right hand dominant. All participants were fluent in the English language and had lived in Australia for over 20 years. However, 30% were bilingual with 85% of bilingual participants having English as a second language. Specifically, 18% spoke Macedonian, 5% spoke French, 5% spoke Italian and 2% spoke Greek. The mean number of years of education was 12.4 (SD = 2.30) with a range of 9 to 18 years. Thirty-six percent of participants had below 12 years of education, 34% had completed 12 years and 30% were educated over 12 years. No participant had previously been administered the Wechsler Memory Scale – Revised, in particular the Visual Reproduction subtest, or had undertaken a neuropsychological assessment in the past.

Due to the well documented effects of depression on cognitive functioning, all participants were screened for depression. According to the Beck Depression Inventory – Second Edition, the mean score for depression for the sample was 5.27 (SD = 4.14) with scores ranging from 0 to 13. All participants except one scored in the minimal range, and as a result only one participant was excluded from this study on the basis of a mood disorder.

Exclusion criteria for the study included a previous history or a current history of a condition where cognitive functioning could significantly improve or deteriorate. This included, but was not limited to the following conditions: a neurological condition or disorder (e.g. brain tumour, stroke, epilepsy, head injury with loss of consciousness greater than five minutes, encephalitis, amnesic syndrome); a progressive or degenerative disease (e.g. multiple sclerosis, dementia, Parkinson’s disease, Huntington’s disease); concurrent drug and alcohol abuse or history of prolonged and excessive alcohol consumption; a history or current treatment of a psychiatric illness (e.g. schizophrenia, bipolar disorder). Participants were not excluded from the study based on risk factors for stroke such as uncontrolled hypertension, diabetes mellitus, or cerebrovascular disease.

2.2. Materials

The Speed and Capacity of Language Processing

(Baddeley, Emslie & Nimmo-Smith, 1992)

The Speed and Capacity of Language Processing test is used to provide information about the efficiency of language comprehension. It is comprised of two tests: 1) the Speed of Comprehension Test, which measures the rate of language processing and 2) Spot-the-Word Vocabulary Test, a silent lexical decision making task comprised of words and non-words.

The Speed and Capacity of Language Processing test reportedly has good reliability estimates of .93. Split-half reliabilities of .84 for the first half and .87 second half were reported in the manual. A moderate association with category generation at .52 and colour naming .56, and a low correlation of .20 with a measure of non-verbal intelligence verified the construct validity of the measure (Baddeley et al, 1992).

Symbol Search subtest (Wechsler Adult Intelligence Scale – Third Edition)

(Wechsler, 1997)

The Symbol Search subtest of the Wechsler Adult Intelligence Scale – Third Edition is a measure of non-verbal information processing speed.

Test-retest reliability coefficients were .77 across all age groups from 16-89 years with stability coefficients ranging .74 to .82. The Symbol Search subtest has a moderate-high correlation with coding at .61 and is moderately associated with other subtests of the Wechsler Adult Intelligence Scale – Third Edition (Wechsler, 1997).

The Wechsler Adult Intelligence Scale – Revised

(Wechsler, 1981; De Lemos, 1981)

The Wechsler Adult Intelligence Scale – Revised is a general measure of intelligence. The Wechsler Adult Intelligence Scale – Revised was chosen instead of the Wechsler Adult Intelligence Scale – Third Edition as it was co-normed with the Wechsler Memory Scale – Revised (see Design section 2.3.1 for further detail).

The Wechsler Adult Intelligence Scale – Revised is comprised of eleven tests: six verbal (Information, Vocabulary, Similarities, Digit Span, Arithmetic and Comprehension), and five non-verbal subtests (Object Assembly, Block Design, Digit Coding, Picture Completion and Picture Arrangement). The subtests are used to derive Verbal and Performance scores and an overall Full Scale IQ score.

An Australian adaptation of the Wechsler Adult Intelligence Scale – Revised is also available. The Australian adaptation substitutes Australian idioms and the content of some of the items in the Information, Arithmetic, Comprehension and Similarities subtests with minor changes to the instructions for the Picture Completion and Block Design subtests also employed (Wechsler, 1981; De Lemos, 1981).

The reliabilities for the Verbal, Performance, and Full Scale IQ indices were reportedly strong with correlation coefficients of .97, .93, .97 respectively. Individual subtest coefficients ranged from .52 for Object Assembly to .96 for Vocabulary with only six of the eleven tests falling below .70. Average reliability coefficients were documented to be generally higher for the Verbal than those of the Performance subtests. The test-retest reliability revealed a practice effect within two to seven weeks between re-administration. The validity of the battery as a measure of general intelligence has been documented with coefficients of .85 with other measures of intelligence (i.e., Stanford-Binet Intelligence Scale) (Wechsler, 1981; Thorndike, Hagen & Sattler, 1986).

The Logical Memory subtest (Wechsler Memory Scale – Revised)

(Wechsler, 1987)

The Logical Memory subtest is a measure of memory for meaningful verbal information in the form of two stories. The Logical Memory subtest has demonstrated good reliability and item validity, consistently loading with other verbal memory tasks and on separate factors from language comprehension measures (Larrabee & Curtiss, 1995).

Rey Auditory Verbal Learning Test

(Rey, 1964; Taylor 1959)

The Rey Auditory Verbal Learning Test is a list-learning task evaluating verbal memory and learning ability. It requires the learning of 15 words, an interference task, immediate recall, delayed recall and recognition procedures. The total score over five learning trials has been identified as the most reliable score at .77. The Rey Auditory Verbal Learning Test has moderate test-retest reliability over a one year interval. Construct validity with verbal memory measures has been reported (Delis 1989; Lezak et al, 2004).

Visual Reproduction subtest (Wechsler Memory Scale – Revised)

(Wechsler, 1987)

The Visual Reproduction subtest is a measure of visual or non-verbal memory. It was described in detail in the previous section 1.12: The Visual Reproduction subtest of the Wechsler Memory Scale: A Closer Look.

Rey Complex Figure Test

(Rey, 1964)

The Rey Complex Figure Test is a measure of non-verbal memory function. The Rey Complex Figure Test involves the reproduction of a complex two-dimensional geometric figure that has a copy, three-minute immediate recall, and a 30-minute delayed recall and recognition procedure. Inter-rater reliability coefficients from .93 to .99 have been documented (Meyers & Meyers, 1995).

Beck Depression Inventory – Second Edition

(Beck, Steer, & Brown, 1996)

The Beck Depression Inventory – Second Edition is a 21 item self-report measure developed for the screening of depressive symptomology that provides only an estimate of the overall severity of depression: 0 – 13 minimal, 14 – 19 mild, 20 – 28 moderate, 29 – 63 severe depression.

The measure has reported high internal consistencies of .92, .93 and .86 and test-retest reliability of .93 (Beck et al, 1996). Construct and concurrent validity has been demonstrated with other measures of depression and anxiety respectively.

Visual Reproduction subtest: Alternative Form

The Alternative Form was developed for the current study and is described in more detail in the Design section.

2.3. Design

2.3.1. Test Selection

A comprehensive neuropsychological evaluation of memory requires the examination of general intellectual ability, speed of information processing, verbal and non-verbal skills and emotional wellbeing, as these factors can affect the efficiency of memory function (Erickson & Scott, 1977; Larrabee & Curtiss, 1995). Therefore, well established and clinically useful measures were administered to ensure useful and accurate interpretations were made on the performance on the Visual Reproduction subtest of the Wechsler Memory Scale – Revised and the Alternative Form. The following established measures were selected:

- The Wechsler Adult Intelligence Scale –Revised was chosen instead of the Wechlser Adult Intelligence Scale – Third Edition, as it was co-normed with the Wechsler Memory Scale – Revised in a large normative study. Therefore, the shared advantage of a common normative group allows for predictions regarding general intellectual ability and memory function to be made.

The prorated short form of the Wechlser Adult Intelligence Scale – Revised (Australian Adaptation) was used to obtain an estimated overall level of intellectual ability. The short form comprises of the following subtests: Information (measure of general knowledge), Vocabulary (understanding the meaning of words), Similarities (ability to make abstract associations between unrelated words or objects), Digit Span (attentional capacity), Arithmetic (mathematical and working memory ability), Picture Completion (identifying essential from non-essential information), Object Assembly (visuo-constructional skills), Block Design (novel problem solving and visuo-constructional ability) and Coding subtests (psychomotor speed). The subtests were administered and scored according to the procedures outlined in the Wechlser Adult Intelligence Scale – Revised manual.

The prorated short form of the Wechlser Adult Intelligence Scale – Revised was used to determine how representative the general abilities of the sample in this study were to the published normative sample. This would also provide useful comparisons to be made between performances on the Visual Reproduction subtest of the Wechsler Memory Scale – Revised and the Alternative Form. The Wechsler Adult Intelligence Scale – Revised was also used to establish the construct validity of the Alternative Form.

- The Speed and Comprehension of Language Processing test and the Symbol Search subtest were used to determine the speed of processing verbal and non-verbal information and whether this confounded performance on the Visual Reproduction subtest of the Wechsler Memory Scale – Revised and the Alternative Form. In addition, these information processing measures were used in a repeated measures design to check the stability of test performances administered over two sessions. Performance on Session One was used as a benchmark in order for Session Two to proceed.

- The Rey Auditory Verbal Learning Test and Logical Memory subtest are verbal memory measures used to evaluate whether discriminant validity of the Alternative Form was established in this sample.

- The delayed recall procedure of the Rey Complex Figure Test was utilised as a measure of nonverbal memory in order to evaluate whether convergent validity of the Alternative Form was established in this sample.

- The Beck Depression Inventory – Second Edition was used as a screening measure to ensure test performances were not confounded by the presence of a depressive mood disorder.

A protocol for the order of test administration was developed to provide a useful and standard process in the administration of the measures between the two sessions. The stability measures would give a benchmark from which the test performances can be interpreted. The memory measures were followed by subtests from the Wechsler Adult Intelligence Scale – Revised to ensure a 25-35 minute time delay between immediate and delayed recall for the memory measures was achieved. The evaluation of depressive symptomology was designed to follow delayed recall on the final session irrespective of the session order the participant was allocated, so if any emotional reactions were provoked by the scale this would not confound test performance.

Verbal and non-verbal measures were alternated in order of administration to maintain motivation and potentially allow successes and failures to be interspersed. This would be particularly relevant for future research in clinical populations where explicit verbal or non-verbal deficits are prevalent (i.e., left and right hemisphere strokes).

The measures used in this study were administered to participants in a serial counterbalanced order between two sessions. This was done so the potential gain from familiarity with the assessment process would be minimised and not confound interpretation of test performance. Session order was determined at the time of recruitment to ensure an equal number of participants were allocated to each session. In the instances when a couple participated in this study, both individuals were allocated the same session that was administered on the same day in order to eliminate the potential for coaching.

Once all measures were administered, the Alternative Scoring System was compared to the Original Scoring System of the Visual Reproduction subtest of the Wechsler Memory Scale – Revised. The relationship between the Alternative Form with other neuropsychological measures was also included in the design of this study to evaluate the quality and the validity of the alternative measure. In particular, to determine what aspects of the Alternative Scoring System were needed to be modified in order to improve the utility of the Alternative Form.

2.3.2. The Development of the Alternative Designs

A previous study conducted by Clark (2000) contributed to the development of this study. Clark (2000) revised the Original Scoring System of the Visual Reproduction subtest of the Wechsler Memory Scale – Revised in an attempt to address its shortcomings (see Appendix A for the Original Scoring System and the Revised Scoring System). This study aimed to develop a scoring system for an alternative form of the Visual Reproduction subtest of the Wechsler Memory Scale – Revised and adapted the scoring principles from Clark (2000) in its development.

In order for the Alternative Form to potentially be interchangeable with the Visual Reproduction subtest of the Wechsler Memory Scale – Revised, the Alternative Designs had to be comparable. As such, a four design structure for the Alternative Form was developed. The Alternative Designs were developed to reflect a similar level of difficulty to the designs of the Visual Reproduction subtest of the Wechsler Memory Scale – Revised.

The following guidelines were adopted:

- Designs A and B of the Alternative Form were developed to be easier than Designs C and D;
- Design A was developed to be more spatially demanding than Design B;
- Design C was developed to be easier than Design D;
- Design C has greater availability to verbal encoding than Designs A, B and D.
- The Alternative Designs would be different from each other so to limit the potential for overlap or confusion between the four designs. That is, each design was carefully devised not to constitute a part of or be closely related to a primary element from another design.

Design A of the Alternative Form was developed to have similar features as Card A of the Visual Reproduction subtest of the Wechsler Memory Scale – Revised. That is, Card A of the Visual Reproduction subtest of the Wechsler Memory Scale – Revised was made up of lines and squares, therefore lines and squares were also used in the development of Design A of the Alternative Form.

Design B was developed to be the most regular in structure and content compared to the other three designs and to represent similar demands as Card B of the Visual Reproduction subtest of the Wechsler Memory Scale – Revised. That is, Card B of the Visual Reproduction subtest of the Wechsler Memory Scale – Revised has circles enclosed within one another and in the same way, Design B of the Alternative Form uses an array of triangular shapes. Namely, both versions use a single regular figure.

Design C was developed to be greater in complexity than Designs A and B but also provide greater availability to verbal encoding. That is, Design C was less spatially demanding than the other three designs, but had a greater number of elements that could easily be verbally labelled. For example, Card C on the Visual Reproduction subtest of the Wechsler Memory Scale – Revised can be interpreted as a window and Design C on the Alternative Form can resemble a pizza.

Design D was developed to be greater in complexity than Designs A, B and C and thus place greater demands on memory. A greater number of different elements of varying spatial complexity were developed for Design D of the Alternative Form to reflect the component structure of Card D of the Visual Reproduction subtest of the Wechsler Memory Scale – Revised, with the development of three primary figures.

Once the four designs were devised, the designs were trialled on individuals with some knowledge and familiarity with neuropsychological test administration. The designs were trialled in order to determine any unrecognised problematic aspects or serious flaws that could be found when administering the design in a test situation. The feedback obtained did lead to some minor refinements to the designs to be made. For example, dots were included within the circles on Design C to increase the complexity of the design. Figure 2.01 shows the four potentially suitable designs used for this study.

Figure 2.01: The Alternative Designs of the Alternative Form

2.3.3. Principles in the Development of the Alternative Scoring System

The development of the scoring system was guided by a number of principles adapted from the previous study conducted by Clark (2000). The principles included the following:

1. Each of the four designs would generate a score out of 20 points. A maximum score of 20 points was chosen for each design (with a total maximum score of 80 points) so that a reasonable range of scores could be obtained and represent a normal distribution of scores if they existed. The score range of the Visual Reproduction subtest of the Wechsler Memory Scale – Revised of 41 points was doubled to 80 points in the Alternative Scoring System, potentially providing a greater range and better distribution of scores.

   In Wechsler Memory Scale - Revised manual there was no rationale for the differential number of points randomly allocated to each of the four designs of the Visual Reproduction subtest. Although it was expected that people will obtain different mean scores on each of the four designs of the Alternative Form, according to the difficulty level and the individual's level of functioning, not assuming a bias on the weighting of each design avoids making arbitrary decisions about the way in which the designs contribute to memory.

2. Furthermore, where multiple figures are present (i.e., Design D), each figure would be given equal weighting and assigned a similar number of scoring criteria. Consequently, no biases or assumptions are made about how the designs are processed and retrieved from memory.

3. The criteria would address as many different aspects of the design rather than a few key elements. This would provide a comprehensive analysis of the designs and potentially a wider range in scores, allow for individuals with poorer reproductions to obtain a score, and circumvent problems with ceiling and floor effects. One of the problems with the Visual Reproduction subtest of the

Wechsler Memory Scale – Revised was that minimal criteria were allocated for each design and limited recall of the design would receive no credit and lead to inaccurate interpretations regarding memory function.

4. Criteria would be scored independently from one another. That is, the score obtained on one item would have no bearing on the score obtained on the following item. For example, Card C on the Original Scoring System states “If criterion 5 is scored zero, criteria 6 through 9 are all scored zero.” Unlike the Original Scoring System, failure on one item on the Alternative Scoring System would not result in failure on the following items. As such no assumptions are made about what is important to be remembered.

5. Despite being verbal in nature, the scoring system would give particular attention to spatial aspects and relationships between elements of the designs to emphasise as much as possible a measure of non-verbal memory.

6. Imperfect recall of the designs would score points. That is, if the reproduction of the design was not perfect, scores could be obtained for the aspects of the design that were accurately recalled. This could minimise floor levels and provide a greater range of scores and better grading of memory performance.

7. Criteria would be allocated for perfect reproductions of the designs. Points would be allocated for accuracy to better differentiate the quality of the reproductions, particularly at the upper end of the score range, so that mildly inaccurate drawings can be discriminated from perfect reproductions.

8. Criteria would be allocated for perfect reproductions that did not contain any extra elements. A penalty for the addition of extra elements or features would be included to better discriminate between good quality drawings and perfect reproductions, thus reducing the potential for ceiling effects.

9. Tolerances for angles, line lengths, minor gaps and curves were included to assist with scoring of the designs, the reliability of scoring and reducing judgement required in interpreting the criteria. Some leniency for the length of lines, precision of angles, and general carelessness in the reproductions would be included so that extraneous factors such as poor motor dexterity or impulsivity did not result in misinterpretations regarding memory ability.

10. A grading of difficulty for each design was developed to assist with the marking process. That is, more complex relationships between elements of the design were assigned criteria later on in the scoring system for the particular design. Tolerances specified for most of the later criteria were included to provide further clarity.

11. The number of criteria that addressed the most frequently reproduced aspects of the design would be limited, so not to make the criteria too simple and have the problem of reaching ceiling levels. Minimising ceiling levels would potentially increase the sensitivity of the measure in detecting mild deficits in non-verbal memory ability.

2.3.4. Five Stages in the Development of the Alternative Scoring System

The development of the Alternative Scoring System occurred in five stages. Stage one involved the generation of a scoring system utilising the scoring principles described in section 2.3.3 to guide its development. Stage two was an initial refinement process that ensured the grading of the difficulty, clarity and meaning of each criterion were adequately represented to minimise misinterpretation. The third stage was a quality checking procedure to ensure the grading of scores was representative of the quality of the reproductions, and identify any obvious problems with the scoring system. Stage four was a secondary refinement process that included the addition of tolerances for criterion and the further revision of the wording and grading difficulty for greater clarity and ease of scoring. The final stage in the development of the Alternative Scoring System included example Full Credit and No Credit drawings to provide a visual guide to further ensure accurate scoring and interpretation. The initial to the final stage in the development of the Alternative Scoring System occurred over a three month period to ensure each stage was developed with a fresh approach. The five stages of development are described in further detail in the Procedure section 2.4.1.

2.3.5. Development of Additional Memory Measures: Cued and Recognition

Additional cueing and recognition procedures were developed to supplement the standard administration of the Alternative Designs. These procedures were developed to provide further clinical and diagnostic information regarding the nature and extent of an underlying memory impairment not provided by the use of free recall alone. Specifically, the cueing and recognition procedures were developed to assist in determining whether memory difficulties were primarily an encoding, storage or retrieval deficit.

Cue and recognition measures can provide information regarding the degree of assistance required to facilitate retrieval of material. Moreover, if some assistance in the retrieval of this information is required, what degree of specification would generate recall of the material: Would a cue be sufficient or would the representation of the entire design be required? In the event that a cue was ineffective in facilitating recall, the use of a recognition procedure would provide information regarding the extent of the difficulty with retrieval and determine whether this material had actually been stored.

The clinical significance in using cue and recognition procedures has been recognised in better understanding the severity of memory problems and assisting with differential diagnosis, which ultimately has important implications for management and rehabilitation.

A shortcoming in the development of the Visual Reproduction subtest of the Wechsler Memory Scale – Revised was the lack of cued and recognition procedures to supplement immediate and delayed recall. Gass (1995) developed a cued procedure, but the un-standardised administration of the cued procedure complicated the use of normative data.

A recognition procedure was developed by Fastenau (1996), but poor psychometric properties rendered it unreliable.

In addition to revising the scoring system of the Visual Reproduction subtest of the Wechsler Memory Scale – Revised, Clark (2000) developed cued and recognition procedures in an attempt to further improve the clinical utility of the measure (see Appendix B). The development of cued and recognition procedures in this study was modelled on the research by Clark (2000) and is described in greater detail in sections 2.3.6 and 2.3.7.

2.3.6. Development of a Cued Procedure for the Alternative Form

A cued recall procedure was developed in this study due to the clinical observation that information previously difficult to recall was facilitated when a cue was provided. That is, a cue may be sufficient to trigger recall without the need for presenting the entire design. A cue procedure can be considered an intermediate step between free recall and recognition. The principles in the development of the cues in this study were adopted from the research by Clark (2000) and included the following three principles:

1. The cue for each design can not be a main feature or contribute to a substantial part of the design. That is, a partial aspect or segment that was representative of the design would be chosen and not an aspect that could be viewed as trivial. For example, using a straight line as a cue for Design A would be insignificant and ineffective as Design B, Design C and Design D also have straight lines included in their design.

2. All the cues would be different and distinguishable from one another. That is, the cue chosen would be a specific aspect of the design that would not be related to any other design, so to eliminate any confusion regarding the design being represented.

3. The cued procedure would follow the standard administration of immediate and delayed recall so performance was not confounded by the provision of a cue. Furthermore, the cue would only be provided if no part of a particular design was generated by free recall. That is, if partial recall of the design was produced, no cue would be provided.

Figure 2.02 shows the cues developed for the Alternative Designs of the Alternative Form.

Figure 2.02: The cues for the Alternative Designs of the Alternative Form.

The cue developed for Design A was a square with a line going through it from the top left to the bottom right corner of it. The square was the best available cue for Design A as a straight line could account for any of the three designs, was less informative and potentially confusing. The use of intercepting lines was considered a substantial feature of the design and determined unsuitable. Furthermore, an empty square could be viewed as obscure, as a square was featured in Design C where a square is formed by the two intersecting lines, and on Design D a square formed the base of the flag figure. The cue developed was congruent with the characteristic item feature of the design and was representative without constituting a major aspect of the design.

The cue for Design B was a triangle. No other figure than a triangle could have been chosen as Design B is made up of an arrangement of triangular figures and logically represented the design.

A circle was chosen to represent Design C. Of the four designs, a circular figure was represented only in Design C and naturally favoured this as a cue for this design.

A number of cues could have been chosen for Design D however a flag was decided upon to best represent the design. Keeping the principles in mind, the alternative choices for the cue for Design D are potentially ambiguous or provide a substantial portion of the total design. For example, if the pack man figure depicted in the middle of Design D was used as a cue, a third of the design would have been revealed. Furthermore, the presentation of a triangle or a square would be confusing as to which design was being represented. A triangle as a cue could only be provided for Design B (as it was a triangular figure) and a square figure could only best represent Design A due to the limited number of elements included of the design.

2.3.7. Development of a Recognition Procedure for the Alternative Form

A recognition procedure was developed in this study to assist in understanding the contribution of retrieval on memory performance. The utility of a recognition procedure in clinical practice is well demonstrated and is supported by current theoretical understanding of the construct of memory. The development of a recognition procedure for the Alternative Form was guided by the research of Clark (2000) that included the following six principles:

1. A 2 x 3 array of six multiple choice items would be devised so to minimise a target decision being made on chance alone. The effects of chance performance are very pronounced in YES/NO recognition formats and are extensively criticised in clinical and research fields alike for the opportunity of chance to guide performance and misrepresent actual memory ability.

2. The target design would be placed amongst a number of distracters on the one layout. Having both the target design and distracters presented to the individual at one time to minimise demands on attention and irrelevant aspects of memory.

3. The five distracter designs would include a common element of each of the target designs and reflect the same number of elements. For example, Design D has three components and so the distracter designs were all developed to contain three figures. However, the distracter designs would also be significantly different from the target design so that a substantial amount of the target design would have to be remembered in order to be adequately recognised. This ensures the decision for each design was based on true recognition of the design and not based on chance.

4. One recognition card would be developed per design. This was devised so to eliminate any additional demands placed on memory and as an efficient way in determining the contribution of retrieval on memory performance. Having one recognition card per design provides a quick means of doing an assessment and reduces the potential for confusion and fatigue in individuals.

5. The recognition procedure would be presented following standardised immediate and delayed recall procedures and only after the cues for the appropriate designs were provided if they were required. Thus the administration process of the recognition procedure was developed so to not contaminate free recall and memory by the provision of recognition cards.

6. The target design was placed in a different relative position on each of the four recognition cards. No target design was placed in the first or last position in the array because the natural bias in guessing tends to favour these positions.

The recognition procedure would also be administered to all participants. That is, even if participants generated a particular design on delayed free recall could the target design be correctly identified from the distracters? This was done in order to determine whether traditional patterns of performances were demonstrated in a non-clinical sample and to evaluate the quality of the procedure. The recognition cards developed for each of the four designs are shown in Figure 2.03.

Figure 2.03: The recognition cards for the Alternative Designs of the Alternative Form

2.3.8. The Development of a Standardised Administration Protocol for the Alternative Form

A standardised administration of the alternative form and the cue and recognition procedures was developed to ensure reliable administration and <u>brief</u> pertinent test instructions. This minimised the demands placed on memory resources so that performance on the task would not be confounded by the exhaustion of other cognitive faculties such as attention. A standardised means of test administration also ensured that performance was not confounded by changes in the wording of instructions and the presentation of material. Figure 2.01 outlines the administration protocol developed in this study.

2.4. Procedure

2.4.1. The Development of the Alternative Scoring System

Stage one in the development process was designed to generate a scoring system taking into consideration the guiding principles outlined in section 2.3.3.

Stage two consisted of an initial refinement process of the scoring system. The 88 protocols obtained from the immediate and delayed performances of the participants in this study were qualitatively assigned to two categories: ‘poor’ and ‘good’ reproductions. The protocols were assigned to categories based on the subjective judgement of the researcher as to whether the quality of the drawing was a ‘poor’ or ‘good’ reproduction of the particular design.

Table 2.01: The administration protocol for the Alternative Designs and the Cued and Recognition measures.

Place a blank piece of paper and a pencil in front of the client……

SAY: I am going to show you some drawings one at a time that I want you to try and remember. I will show you the drawing for 10 seconds and when I take it away I want you to draw it. Make a copy of it when I take it away. Ready?

Show Design A for 10 seconds……Take drawing away

Allow client to finish……

SAY: Have a look at this drawing for 10 seconds. Make a copy of it when I take it away.

Show Design B for 10 seconds……Take drawing away

Allow client to finish……

SAY: Now look at this drawing for another 10 seconds.

Show Design C for 10 seconds……Take drawing away

Allow client to finish……

SAY: The next drawing has more than one part to it. Look at it for 10 seconds and draw what you can remember. Make a copy of it when I take it away.

Show Design D for 10 seconds……Take drawing away

Allow client to finish……

After a time delay of 25-35 minutes…..

Place a blank piece of paper and a pencil in front of the client……

SAY: Remember the drawings that I showed you before. Can you draw what you remember of them on now?

If client states that they cannot recall any of them……

SAY: Have a bit more of a think about it....There were four drawings I showed you… can you draw for me what you can remember even if it is only a little part.

Administration of Cue Measures

If client completely fails to recall any one of the drawings, or no part of their drawing is recognisable about which design it represents, then draw the standard partial cue for the appropriate design

SAY: Does this help?

Allow client to finish……

Administration of Recognition Measures

Show each recognition card for Designs A – D one at a time. Place the recognition card for Design A it in front of client……

SAY: Is any one of these drawings here the one I showed you before?

If client makes more than one selection get them to choose only one.

Repeat the above for Designs B, C and D.

A subset of 48 protocols consisting of both ‘poor’ and ‘good’ immediate and delayed recall designs were selected based on the researcher’s view of design quality to assist in the difficulty grading of the item and refinement of the scoring system. The six poorest reproductions for each of the four designs were utilised to ensure the scoring system assigned items to address all the aspects of the design and to have a score that was representative of the quality of the reproduction. Six good reproductions of varying accuracy for each of the four designs were utilised to ensure that criteria would be able to differentiate between the good, better and perfect representations of the designs and adequately represent the quality of the design reproduced.

The initial refinement process also involved the scrutiny of the scoring criteria for any ambiguous, confusing or duplications in the wording of all scoring item for each of the designs.

Stage three was a quality checking procedure to make sure the scoring system was not biased and reflected the impression given by looking at the protocols. This stage was not a sophisticated means of analysis rather a check to determine that no obvious faults were inherent in the scoring criteria for each design.

Stage four was the final refinement stage in the development of the scoring system. In this stage tolerances were included for the length of lines, angles and carelessness in reproductions. Exact tolerance ranges were developed so that scoring criteria were not open to interpretation and judgements would not have to be made as to whether an item was met. For example, a tolerance of 20 degrees was provided between the interception between the horizontal and vertical lines on Design B, so an angle measuring between 80 and 100 degrees would be acceptable and score credit for the item.

In the secondary refinement process the scoring system was again analysed to ensure criteria were clearly described so no ambiguity regarding the aspect of the design that was being addressed. Criteria that were harder to score were modified to eliminate the potential for differing interpretations. In addition, tolerances were included particularly for the later criteria where the complexity of the relationships between elements increased in order to make criteria more explicit. For example, each of the three primary figures of Design D was segregated into sections to better represent the aspects of the elements of the design that were being addressed by the criteria.

In the final stage of development, examples of no credit and full credit drawings were developed. The addition of example drawings was devised to complement the verbal nature of scoring criteria with a visual means of representing the item to provide greater clarity and to assist with interpretation of each design. The example drawings included actual protocols obtained from the participants in this study and potential misinterpretations of the designs considered by the researcher.

2.4.2. Development of Cued and Recognition Procedures

This study also involved the development of a cueing procedure and a multiple choice recognition format for the Alternative Form. The cues for the Alternative Designs were developed by reviewing each of the designs carefully and identifying a characteristic feature of the design that was unique to and representative of the individual design. When the four cues for each design were decided upon the cues were checked for any ambiguity or misinterpretations of which design was being represented.

The recognition cards for the Alternative Designs were developed after careful consideration of the potential misinterpretations of the elements of each design. The development of the distracter designs was devised to retain a similar level of complexity as the target design and a similar number and type of elements. That is, the angles, rotations, order and position of the elements in relation to one another were altered and re-arranged in an attempt to make them distinctly different so not to result in confusion. Conversely, the distracter designs were developed to be similar in shape and size. The recognition cards were developed one at a time over a month period. After a two week period minor refinements were made to the distracter designs. For example, additional closed circles were included for first, third and sixth distracters for Design C to qualitatively appear similar in the level of complexity.

2.4.3. The Administration of the Alternative Form in a Non-Clinical Sample

Ethical approval for the study was obtained from the Victoria University Ethics Committee and all procedures were carried out in accordance with the university and standard ethical guidelines. All participants were initially contacted by telephone and the purpose of the study and procedures were outlined. Once verbal consent was obtained, a convenient time and day for the assessment was arranged in the comfort of the participant's home. Upon arrival to their homes, participants were provided with the information sheet outlining the study (see Appendix B) and were encouraged to ask questions on interview. Prior to the commencement of testing, participants were invited to sign a consent form indicating voluntary consent to participate (see Appendix C). Participants were advised withdrawal from the study could occur at any time without any obligations or personal consequences.

As the research involved two forms of the Visual Reproduction subtest (the Wechsler Memory Scale – Revised version and the Alternative Form developed in this study), the measures were administered over two sessions scheduled six to eight days apart.

Session one involved obtaining demographic and background information (occupation, age, gender, ethnicity, educational history, medical and psychiatric history) from each participant. English language proficiency was established throughout the interview process prior to the standardised administration of the measures.

In both sessions one and two the stability measures, namely the Symbol Search subtest and the Speed and Capacity of Language Processing test were initially administered. The administration of the memory measures followed which included the Logical Memory subtest and the Visual Reproduction subtest of the Wechsler Memory Scale – Revised in the first session; and the Alternative Form, the Rey Auditory Verbal Learning Test and the Rey Complex Figure Test in the second session. The Wechsler Adult Intelligence Scale – Revised subtests were administered in verbal and non-verbal succession. The Beck Depression Inventory – Second Edition was administered at the end of the final session irrespective of the session order that the participant was allocated. The order of test administration is outlined below in Table 2.02.

Table 2.02: The order of test administration for Session One and Session Two.

Session One
1) Demographic and medical history
2) Symbol Search subtest of the Wechsler Adult Intelligence Scale – 3rd Edition (SS)
3) Speed and Capacity of Language Processing (SCOLP)

4) Logical Memory subtest (WMS-R)
5) Visual Reproduction subtest (WMS-R)
6) The Wechsler Adult Intelligence Scale –Revised:
Digit Span, Picture Completion, Vocabulary, Coding, Similarities subtests

7) Delayed recall of Logical Memory
8) Delayed recall of Visual Reproduction

Session Two
1) Symbol Search subtest of the Wechsler Adult Intelligence Scale – 3rd Edition (SS)
2) Speed and Capacity of Language Processing (SCOLP)

3) Visual Reproduction Alternative Designs (VRalt)
4) The Rey Auditory Verbal Learning Test (RAVLT)
5) Rey Complex Figure Test (RCFT)

6) The Wechsler Adult Intelligence Scale –Revised:
Arithmetic, Block Design, Information, Object Assembly subtests

7) Delayed recall and recognition of Visual Reproduction Alternative Designs (VRalt)
8) Delayed recall and recognition of RAVLT
9) Delayed recall of Complex Figure of Rey

10) Beck Depression Inventory (BDI-II)

On completion of the two sessions, participants were offered a feedback session and a suitable date and time was arranged. Approximately two weeks after testing was completed participants received verbal feedback via telephone with a general summary of their performance.

RESULTS

3.1. The Alternative Form

3.1.1. The Alternative Designs

A primary aim of the study was to develop a scoring system for an alternative form of the Visual Reproduction subtest of the Wechsler Memory Scale - Revised. In order to develop a scoring system for the Alternative Form, four Alternative Designs were required to be initially developed. The Alternative Designs developed in this study are shown in Figure 3.01 and described in detail in the previous Design section 2.3.2.

Figure 3.01: The Alternative Designs for the Alternative Form

3.1.2. Additional Memory Measures: Cue and Recognition Procedures

In addition to the development of Alternative Designs, cue and recognition procedures were devised to complement the Alternative Form to obtain additional potentially clinically useful information about memory function. Figure 3.02 and Figure 3.03 illustrate the cued and recognition procedures developed in this study that are described in detail in the previous Design sections 2.3.6 and 2.3.7 and Procedure section 2.4.2.

Figure 3.02: The Cues for the Alternative Designs of the Alternative Form

Figure 3.03: The Recognition cards for the Alternative Designs of the Alternative Form

3.1.3. The Alternative Scoring System

The Alternative Scoring System for the Alternative Form of the Visual Reproduction subtest of the Wechsler Memory Scale – Revised was developed utilising the principles outlined in the Design section 2.3.3. The scoring system derived for each of the four designs is presented below in Table 3.01 for Design A; Table 3.02 for Design B; Table 3.03 for Design C; and Table 3.04 for Design D. A general scoring rule for each design was specified to independently score each item, so that failure on one would not imply failure on another item. Each item would be scored one point or zero points depending on whether that item was represented in the recall of the design with a maximum total score of 20 points awarded. Examples of full credit drawings and no credit drawings were developed for each item to provide further clarity and assist with the interpretation of the items of the scoring systems for each of the four designs.

Table 3.01: The scoring system for Design A of the Alternative Form

20 items **Each item is worth one point** **Score each item independently**	
Item 1:	At least two continuous lines are present that are no shorter than 2cm in length. If there is only one line, score zero. Tolerance: Mild curves in lines are not penalised. Example: Full Credit Drawing Example: No Credit Drawing
Item 2:	There are only two continuous lines present. One of these lines is drawn diagonally from the lower left to the upper right quadrant and is at least 2cm in length. Tolerance: Mild curves in lines are not penalised. Example: Full Credit Drawing Example: No Credit Drawing
Item 3:	There are at least two continuous lines present and these two lines touch or intersect (or four lines that emanate from a central point). Lines are at least 2cm in length. Tolerance: Mild curves in lines are not penalised. Example: Full Credit Drawing Example: No Credit Drawing
Item 4:	At least one solid figure is clearly present (e.g. a square, triangle or even a circle etc). If there is no solid figure, score zero. Tolerance: Mild curves in lines and gaps are not penalised. Example: Full Credit Drawing Example: No Credit Drawing
Item 5:	At least two separate solid figures are clearly present and one is a quadrilateral (e.g. a four-sided figure: a square, diamond, rectangle etc). Tolerance: Mild curves and gaps are not penalised. Example: Full Credit Drawing Example: No Credit Drawing

Table 3.01 continued….

Item 6: Only two separate solid figures are clearly present and both are quadrilaterals.
Tolerance: Mild curves and gaps are not penalised.

Example: Full Credit Drawing Example: No Credit Drawing

Item 7: There is at least one continuous line (not shorter than 2cm in length) located between any two separate figures so that it forms a link between them. Figures do not have to be quadrilaterals.
Tolerance: Line(s) does not have to touch either figure and can intersect (no more than 10% of the figure of the length of the line).

Example: Full Credit Drawing Example: No Credit Drawing

Item 8: There are only two continuous lines (not shorter than 2cm in length) located between any two separate figures so that it forms a link between them. Figures do not have to be quadrilaterals.
Tolerance: Lines do not have to touch either figure and can intersect (no more than 10% of the figure or the length of line).

Example: Full Credit Drawing Example: No Credit Drawing

Item 9: There are only two separate figures that are located relative to each other with the top of the right figure clearly above the top of the left figure. One figure is positioned to the lower left of the other. Figures do not have to be quadrilaterals.
Tolerance: No more than 25% overlap of the vertical dimension of the figures is accepted.

Example: Full Credit Drawing Example: No Credit Drawing

Item 10: There are only two continuous lines that only touch two separate figures that are both quadrilaterals.
Tolerance: Minor overshoots, gaps and mild curves in lines are not penalised.

Example: Full Credit Drawing Example: No Credit Drawing

Table 3.01 continued….

Item 11:	Nominate the left quadrilateral figure as (A) and the right quadrilateral figure as (B). One straight continuous line goes from the top or bottom right corner of the left quadrilateral (A) to touch the other right quadrilateral (B) anywhere at any point. Tolerance: Position – Line must emanate from 1/5 of the length of the top or bottom right side of quadrilateral (A). Tolerance: Minor overshoots, gaps and mild curves in lines are not penalised. Example: Full Credit Drawing Example: No Credit Drawing
Item 12:	One continuous line goes from the top or bottom left corner of the right quadrilateral (B), to touch the left quadrilateral (A) anywhere at any point. Tolerance: Position – Line must emanate from 1/5 of the length of the top or bottom left side of quadrilateral (B). Tolerance: Minor overshoots, gaps and mild curves in lines are not penalised. Example: Full Credit Drawing Example: No Credit Drawing
Item 13:	One continuous line is drawn diagonally from the bottom right corner of (A) to the top left corner of (B). Tolerance: Position – Line must emanate from the bottom 1/5 of the length of the right side of quadrilateral (A) and end within the top 1/5 of the left side of quadrilateral (B). Tolerance: Minor overshoots or gaps and mild curves in lines are not penalised. Example: Full Credit Drawing Example: No Credit Drawing
Item 14:	One continuous line is drawn diagonally from the bottom left corner of (B) to the top right corner of (A). Tolerance: Position – Line must emanate from the bottom 1/5 of the length of the left side of quadrilateral (B) and end within the top 1/5 of the left side of quadrilateral (A). Tolerance: Minor overshoots or gaps and mild curves in lines are not penalised. Example: Full Credit Drawing Example: No Credit Drawing
Item 15:	At least one small line is present within each quadrilateral. Tolerance: Line(s) do not have to exactly touch and can overshoot a little. Example: Full Credit Drawing Example: No Credit Drawing

Table 3.01 continued….

Item 16:	Only one diagonal line is present in each quadrilateral and it touches the top left and runs through to the bottom right corner of each quadrilateral. Tolerance: Minor overshoots and gaps are accepted. Example: Full Credit Drawing Example: No Credit Drawing
Item 17:	The two quadrilateral figures present are both squares. Tolerance: Angle – Each Square can form an angle no less than 70 degrees and no more than 110 degrees. Tolerance: Length – The length of the sides of the square are no shorter than 1/5 of the longest side. Tolerance: Mild curves, gaps or overshoots are not penalised. Example: Full Credit Drawing Example: No Credit Drawing
Item 18:	One major continuous line is close to horizontal. The length of this horizontal line, which is present between the two squares, is approximately two to three times the length of the bottom side of the right square. Tolerance: Line is not rotated more than 10 degrees from horizontal plane. Tolerance: Mild curves in line are not penalised. Example: Full Credit Drawing Example: No Credit Drawing
Item 19:	Only two major continuous lines intersect and form an angle no less than 30 degrees and no more than 45 degrees. Example: Full Credit Drawing Example: No Credit Drawing
Item 20:	No extra elements are present, according to the model. Tolerance: Minor overshoots of lines and crossed out elements are not penalised. Example: Full Credit Drawing Example: No Credit Drawing

Table 3.02: The scoring system for Design B of the Alternative Form

20 items
Each item is worth one point
Score each item independently

Item 1: A triangular figure is present.
Tolerance: Minor curves or gaps in lines are not penalised.

Example: Full Credit Drawing Example: No Credit Drawing

Item 2: More than one triangular figure is clearly present.
Tolerance: Minor curves or gaps in lines are not penalised.

Example: Full Credit Drawing Example: No Credit Drawing

Item 3: At least two horizontal lines are present. They may be part of or contiguous with the triangles i.e. run parallel.
Tolerance: Minor curves or gaps in lines are not penalised.

Example: Full Credit Drawing Example: No Credit Drawing

Item 4: Only two continuous horizontal lines are present and they are of equal length.
Tolerance: The smaller line is at least 80% of longest line.
Tolerance: Minor curves in lines are not penalised.

Example: Full Credit Drawing Example: No Credit Drawing

Item 5: At least one vertical line is present.
Tolerance: Minor curves in line(s) are not penalised.

Example: Full Credit Drawing Example: No Credit Drawing

Item 6: Only one continuous vertical line is drawn that is 50-100% length of longest horizontal line.
Tolerance: Minor curves in line are not penalised.

Example: Full Credit Drawing Example: No Credit Drawing

Table 3.02 continued….

Item 7: The Vertical line forms a right angle with at least one horizontal line. An angle smaller than 80 degrees or larger than 100 degrees receives Example: No Credit Drawing, score zero.
Tolerance: Minor curves in lines are not penalised.

Example: Full Credit Drawing Example: No Credit Drawing

Item 8: The Vertical line forms a right angle with two horizontal lines and only two.
Tolerance: Minor curves in lines are not penalised.

Example: Full Credit Drawing Example: No Credit Drawing

Item 9: At least one diamond type figure is clearly present between two continuous horizontal lines. More than one diamond figure can be present.
Tolerance: Minor curves or gaps in lines are not penalised.

Example: Full Credit Drawing Example: No Credit Drawing

Item 10: At least four small triangles are present and are in a mirror image formation.
Tolerance: Minor curves or gaps in lines are not penalised.

Example: Full Credit Drawing Example: No Credit Drawing

Item 11: There are only four small triangles in a mirror image formation that are positioned between two horizontal lines and touch in the middle section i.e the apex of the triangles touch approximately at the half-way point between the average distance from one horizontal line to the other.
Tolerance: Minor curves or gaps in lines are not penalised.

Example: Full Credit Drawing Example: No Credit Drawing

Table 3.02 continued….

Item 12:	Four small triangles are relatively equal in size. Tolerance: Smallest triangle is at least 90% the size of largest small triangle. Tolerance: Minor curves or gaps in lines are not penalised. Example: Full Credit Drawing Example: No Credit Drawing
Item 13:	One large triangle has the horizontal line as its upper side and it is pointing down. There is no need for overlap with another triangle (if one exists), and the presence of only one triangle is acceptable. Tolerance: Minor curves or gaps in lines are not penalised. Example: Full Credit Drawing Example: No Credit Drawing
Item 14:	One large triangle has the horizontal line as its lower side and it is pointing up. There is no need for overlap with another triangle (if one exists) and the presence of only one triangle is acceptable. Tolerance: Minor curves or gaps in lines are not penalised. Example: Full Credit Drawing Example: No Credit Drawing
Item 15:	Only two large triangles are present and each has a horizontal line as its base. The base of one triangle is a horizontal like and this base touches the tip (or apex) of the other triangle and vice versa. Tolerance: Minor overlap and overshoots are accepted. Example: Full Credit Drawing Example: No Credit Drawing
Item 16:	The two large triangles that are formed by the two horizontal lines are of similar size. Tolerance: The smaller large triangle is at least 75% the size of the other large triangle Tolerance: Minor overlap and overshoots are accepted. Example: Full Credit Drawing Example: No Credit Drawing

Table 3.02 continued….

Item 17: The two large triangles form equilateral triangles.
<u>Tolerance:</u> Shortest side of triangle is at least 90% the length of the longest side.
<u>Tolerance:</u> Minor curves, gaps or overshoots in lines are accepted.

<u>Example</u>: Full Credit Drawing <u>Example</u>: No Credit Drawing

Item 18: At least one vertical line is present within at least one large triangle that extends from the base to the apex of the triangle.
<u>Tolerance:</u> 10% overshoot is allowed.
<u>Tolerance:</u> Minor curves in lines are not penalised.

<u>Example</u>: Full Credit Drawing <u>Example</u>: No Credit Drawing

Item 19: Only one vertical line joins two major horizontal lines and is located centrally within both major triangles, according to the model. No overshoots or minor overlap accepted.

<u>Example</u>: Full Credit Drawing <u>Example</u>: No Credit Drawing

Item 20: No extra elements are present.
<u>Tolerance:</u> Minor curves in lines and crossed out elements are not penalised.

<u>Example</u>: Full Credit Drawing <u>Example</u>: No Credit Drawing

Table 3.03: The scoring system for Design C of the Alternative Form

20 items
Each item is worth one point
Score each item independently

Item 1: At least one large figure is present that encloses one or more other elements. The figure can be a quadrilateral, circle, triangle etc.
Tolerance: Mild curves or gaps in lines are accepted.

Example: Full Credit Drawing Example: No Credit Drawing

Item 2: The large figure is circular – can be ovoid but must have no straight edges. If more than one large figure, choose the figure that will maximise total score.
Tolerance: Minor gaps or overshoots are not penalised.

Example: Full Credit Drawing Example: No Credit Drawing

Item 3: There are between two and four internal lines present that divide the large figure and run from one side to the other). If no large figure is present still give credit for two to four lines.
Tolerance: Minor gaps or overshoots are not penalised.

Example: Full Credit Drawing Example: No Credit Drawing

Item 4: Two vertical lines are present that run approximately parallel to one another. There may be a gap at a central intersecting point but both vertical lines need to transverse most of the large figure.
Tolerance: Mild curves in lines are not penalised.

Example: Full Credit Drawing Example: No Credit Drawing

Item 5: Two lines are present that are approximately parallel. There may be a gap at a central intersecting point but both horizontal lines need to transverse most of the large figure.
Tolerance: Mild curves in lines are not penalised.

Example: Full Credit Drawing Example: No Credit Drawing

Table 3.03 continued….

Item 6: Two sets of two parallel lines intersect at right angles. If lines are rotated nominate one as ‘vertical’ and the other ‘horizontal.’ An angle smaller than 80 degrees or larger than 100 degrees receives no credit, score zero.

Tolerance: Minor curves, gaps or overshoots are not penalised.

Example: Full Credit Drawing Example: No Credit Drawing

Item 7: Gap between two vertical parallel lines, or nominated vertical lines, must be within 5-20% of width of large figure.

Tolerance: Mild curves in lines are not penalised.

Example: Full Credit Drawing Example: No Credit Drawing

Item 8: Gap between two horizontal parallel lines, or nominated horizontal lines, must be within 5-20% of width of large figure.

Tolerance: Mild curves in lines are not penalised.

Example: Full Credit Drawing Example: No Credit Drawing

Item 9: The central square typically formed by the two sets of intersecting lines, forms right angles.

Tolerance: Angle: Angles need to be between 80 and 100 degrees with no gaps present.

The lengths of each sides of the square are equal.

Tolerance: Length: The shortest line must be at least 80% of longest. The square needs to be located in the centre of the figure.

Tolerance: Position: Square needs to within the central 20% of the figure. An isolated square is accepted if it meets angle length and position tolerance.

Example: Full Credit Drawing Example: No Credit Drawing

Item 10: At least two dots or small circles are present between the vertical lines or the nominated vertical lines. If dots or small circles lie beside or run parallel to a single vertical line, give credit.

Example: Full Credit Drawing Example: No Credit Drawing

Table 3.03 continued….

Item 11:	At least two dots or small circles are present between the horizontal lines or the nominated horizontal lines. If dots or small circles lie beside or run parallel to a single horizontal line, give credit. Example: Full Credit Drawing Example: No Credit Drawing
Item 12:	There is the same number of dots or small circles in the top half as the bottom half of the figure, if it was segregated at an approximate half-way or central point. Example: Full Credit Drawing Example: No Credit Drawing
Item 13:	There is the same number of dots or small circles in left half as in right half of the figure, if it was segregated at an approximate half-way or central point. Example: Full Credit Drawing Example: No Credit Drawing
Item 14:	Exactly four dots, or filled in small circles, are present in each quadrant segregated by two sets of parallel lines that intersect. Dots cannot be open circles. Example: Full Credit Drawing Example: No Credit Drawing
Item 15:	Four quadrants have at least one small figure within each. Small figure can be any figure e.g. a triangle, square etc. Tolerance: Small figures that touch or mildly overlap with any line or part of the large figure are not penalised. Example: Full Credit Drawing Example: No Credit Drawing

Table 3.03 continued….

Item 16:	Only four small figures are present and there is only one in each quadrant. Small figures can be any figure e.g. a triangle, square, rectangle etc. Tolerance: Small figures that touch or mildly overlap with any line or part of the large figure are not penalised. Example: Full Credit Drawing Example: No Credit Drawing
Item 17:	Only four small figures are present, one in each quadrant, and the small figures are circular (i.e can be ovoid etc). Tolerance: Small figures that touch or mildly overlap with any line or part of the large figure are not penalised. Example: Full Credit Drawing Example: No Credit Drawing
Item 18:	At least two of the four small figures have an internal figure clearly inside it. Small figures can be any figure e.g. triangle, square etc. Dots are accepted. Example: Full Credit Drawing Example: No Credit Drawing
Item 19:	There are only four small figures in each quadrant and this small figure is a circle. There is only one internal figure present in each small figure and this is an open circle. No dots are allowed. Example: Full Credit Drawing Example: No Credit Drawing
Item 20:	No other extra elements are present i.e. dots, lines or figures. Crossed out elements are not penalised. Figure cannot be rotated, according to the model. Example: Full Credit Drawing Example: No Credit Drawing

Table 3.04: The scoring system for Design D of the Alternative Form

20 items
Each item is worth one point
Score each item independently

Section **L** (rectangular figure -left)

Item 1: A figure or a combination of figures forms a square like or rectangular gestalt.
Tolerance: Mild curves in lines are not penalised.

Example: Full Credit Drawing   Example: No Credit Drawing

Item 2: A rectangular or square like configuration has only three elements to it. Or, there are three parts in close proximity to each other on the left hand side that are clearly not part of any other figure present.
Tolerance: Minor overlap between elements should not be penalised.

Example: Full Credit Drawing   Example: No Credit Drawing

Item 3: One of the three figures on the left hand side is a large triangle or a trapezoid. It forms 25- 50% of the surface area of the set of elements in criteria 2.

Example: Full Credit Drawing   Example: No Credit Drawing

Item 4: The large figure is a triangle or a trapezoid and is resting on its vertex.

Example: Full Credit Drawing   Example: No Credit Drawing

Item 5: The large figure is a triangle and has only one right angle.
Tolerance - Angle: Angle needs to be between 80 and 100 degrees with no gaps present.

Example: Full Credit Drawing   Example: No Credit Drawing

Item 6: At least one smaller triangle lies adjacent to the larger triangle.
Tolerance: Minor overlap between the two figures is not penalised.

Example: Full Credit Drawing   Example: No Credit Drawing

Table 3.04 continued….

Item 7: A trapezoid figure lies to the left of the large triangle and is sitting underneath part of it.
Tolerance: Minor overlap between the two figures is not penalised.

Example: Full Credit Drawing Example: No Credit Drawing

Item 8: A trapezoid figure and only one small triangle are positioned to the left of the large triangle and form a triangular gestalt to complement the larger triangle with a clear gap between them.
Tolerance: At least a one mm gap must be present between them.

Example: Full Credit Drawing Example: No Credit Drawing

Section **M** (circular figure -middle)

Item 9: A figure with a curved surface is present and is clearly separate from figures in criteria 2-8.

Example: Full Credit Drawing Example: No Credit Drawing

Item 10: Figure is not a complete circle but it is larger than a semi-circle.

Example: Full Credit Drawing Example: No Credit Drawing

Item 11: A missing segment from the circular figure is 20-30% of the total figure.

Example: Full Credit Drawing Example: No Credit Drawing

Item 12: The missing segment has one right angle.
Tolerance - Angle: The angle formed by the two lines needs to be between 80 and 100 degrees with no gaps present.

Example: Full Credit Drawing Example: No Credit Drawing

Item 13: Missing segment is in upper right quadrant of the circular figure, if circular figure was segregated into four quadrants.

Example: Full Credit Drawing Example: No Credit Drawing

Table 3.04 continued….

Section **R** (Flag – right)

Item 14: At least one line or figure is present. This line or figure is separate and distinct from segments M or L.

Example: Full Credit Drawing Example: No Credit Drawing

Item 15: There is at least one large line with at least one figure attached to it.

Example: Full Credit Drawing Example: No Credit Drawing

Item 16: There is only one large vertical line and at least two attached figures are present. One of these figures is a square or a triangle.

Tolerance - Angle: The angle of the vertical line must be between 80 and 100 degrees.

Example: Full Credit Drawing Example: No Credit Drawing

Item 17: There are only two attached figures to the large vertical line. One figure is a triangle that is positioned towards the top and the other is a square positioned towards the bottom of the vertical line.

Tolerance - Angle: The angle of the vertical line must be between 80 and 100 degrees.

Example: Full Credit Drawing Example: No Credit Drawing

Item 18: The triangle is clearly positioned towards the top on the left hand side of the vertical line and points left, forming a right angle with the vertical line. The square is clearly positioned towards the bottom on the right hand side of the vertical line.

Tolerance – Angle: Angle of the triangle must be between 80 and 100 degrees

Tolerance - Length: Attached figures must not take up more than a third of the length of the vertical line.

Example: Full Credit Drawing Example: No Credit Drawing

Table 3.04 continued….

Item 19: All three major sections are present and elements are located correctly i.e. they are in The right order (rectangular figure (section L), circular figure (section M), flag (section R)), and the bases of all these elements are similar in height. The top of section M is clearly below the tops of sections L and R, according to the model.

Example: Full Credit Drawing Example: No Credit Drawing

Item 20: No extra elements or figures are present. Crossed out elements are not penalised.

Example: Full Credit Drawing Example: No Credit Drawing

3.2. Analysis

Once all raw data was obtained and scored, data were analysed using the Statistical Package for the Social Sciences versions 14.0 and 15.0. Frequency analyses, measures of central tendency (the average scores), standard deviations (measure of the dispersion of scores between one another) the range and distribution of scores were conducted for all variables to provide the context for interpretations to be made.

3.3. Performance Characteristics of the Participants in this Study

3.3.1. Level of Intellectual Functioning

As the Visual Reproduction subtest of the Wechsler Memory Scale – Revised was co-normed with the Wechsler Adult Intelligence Scale – Revised, the prorated short form of the measure was used in this study to obtain an estimate of general intellectual functioning. A Full Scale Intelligence Quotient score and Separate Verbal Intelligence Quotient and Performance Intelligence Quotient scores were prorated according to the normative data provided in the manual (Wechsler, 1987).

The range of ages in this study was 25 to 51 years. The mean age of the sample was 37.91 years, *SD* = 7.60, and fell within the 35 – 44 year age band according to the Wechsler Memory Scale – Revised and Wechsler Adult Intelligence Scale – Revised age bands. Subsequently, the Full Scale IQ mean score for the 35 – 44 age group of 104, *SD* = 17, was used as the reference point for making comparisons in this study. The mean IQ scores and range of ability of the sample are shown in Table 3.05.

Table 3.05: The mean Full Scale IQ, Verbal IQ and Performance IQ score on the Wechsler Adult Intelligence Scale – Revised (n = 44).

WAIS-R	*M*	*SD*	Range
Full Scale IQ	113.0	13.7	89 – 141
Verbal IQ	109.0	12.7	88 – 137
Performance IQ	115.6	14.0	85 – 144

The results revealed that the mean Full Scale IQ score of participants in this study was within the High Average range, 113, *SD* = 13.7, with IQ scores from the Low Average to Very Superior range and above (89 – 141). The mean IQ score of this study was greater than the mean Full Scale IQ reported in the Wechsler Memory Scale – Revised manual for the standardisation sample of 104, *SD* = 17.

The average Verbal IQ score was within the Average range 109, *SD* = 12.7, and Performance IQ score was within the High Average range 115, *SD* =14, with participants demonstrating a relative strength in non-verbal abilities. However, the mean discrepancy between verbal and non-verbal skills of approximately six points was not a large difference and commonly occurs in the general population.

3.3.1.1. Verbal IQ and Performance IQ subtest performances

Table 3.06 shows that on the Verbal Scale of the Wechsler Adult Intelligence Scale – Revised, performances were generally within the upper end of the Average range. Although a wide range of performances were obtained 95% of scores were at or above the Low Average range.

Similarly, scores on the Performance Scale of the Wechsler Adult Intelligence Scale – Revised were generally within the upper end of the Average range. Although a wide range of performances were obtained 91% of scores were at or above the Low Average range, with only one participant having a below average score in visuo-constructional ability.

Table 3.06: Participant scaled score performances on the subtests of Verbal IQ Scale and the Performance IQ Scale of the Wechsler Adult Intelligence Scale – Revised (n = 44).

WAIS-R Subtests	*M*	*SD*	Range
Verbal IQ Scale			
Information	11.2	.39	7 – 18
Vocabulary	11.8	.39	7 – 17
Similarities	11.9	.37	8 – 18
Digit Span	11.0	.37	6 – 18
Arithmetic	11.1	.40	6 – 17
Performance IQ Scale			
Picture Completion	11.6	.33	6 – 17
Object Assembly	11.9	.39	5 – 18
Block Design	12.3	.43	7 – 18
Coding	11.9	.30	6 – 15

3.3.2. Performance on measures of verbal memory and learning

Measures of verbal memory and learning were utilised in this study to determine the discriminant validity of the Alternative Scoring System. The Logical Memory subtest was administered and scored according to the protocol outlined in the Wechsler Memory Scale – Revised manual. The Rey Auditory Verbal Learning Test was administered and scored according to instructions and the normative data published in Spreen and Strauss (1998). The performance on the Logical Memory subtest and the Rey Auditory Verbal Learning Test are shown in Table 3.07 and Table 3.08 respectively.

3.3.2.1. Logical Memory subtest

Table 3.07 shows that after an initial presentation of the Logical Memory stories, the mean score was 28.8, *SD* = .81, which was at the 59th percentile. Although a wide distribution of performances were obtained 13 – 42, most participants scored above the 25th percentile with only one participant scoring below the 25th percentile.

Table 3.07: The mean, standard deviation and range of performances on the Logical Memory subtest (n = 44).

Logical Memory	*M*	*SD*	Range
Immediate Recall	28.8	.81	13 – 42
Delayed Recall	24.3	.97	8 – 36

The mean score following a time delay was 24.3, *SD* = .97, also at the 59th percentile. Although a wide range of scores were obtained 8 – 36, the majority of participants performed well with only four participants performing below the 25th percentile.

3.3.2.2. Rey Auditory Verbal Learning Test

Performance on the Rey Auditory Verbal Learning Test after five presentations provided a mean learning score of 51.5, *SD* = 1.22, that was within age expectations (see Table 3.08). The range of total learning scores was 37 – 72, with most of the scores falling within two standard deviations of the mean, and only three participants performed below age expectations.

The mean score on the Rey Auditory Verbal Learning Test after a delay was 10.7, *SD* = .4, and within age expectations. The range of scores was 5 – 15 with the majority of participants performing within two standard deviations of the mean and only two participants scored below age expectations.

Table 3.08: The mean, standard deviation, and the range of performances on the Rey Auditory Verbal Learning Test (n = 44).

Rey Auditory Verbal Learning Test	*M*	*SD*	Range
Learning Score	51.5	1.22	37 – 72
Delayed Recall	10.7	.40	5 – 15

3.3.3. Measures of Non-Verbal Memory

3.3.3.1. The Rey Complex Figure Test

The Rey Complex Figure Test was used in this study to determine the convergent validity of the Alternative Scoring System. The Rey Complex Figure Test was administered and scored according to instructions and the normative data published in Spreen and Strauss (1998).

Table 3.09 shows that on the copy procedure of the Rey Complex Figure Test the mean score was 33.8, *SD* = .48, which was within the upper end of age expectations. The range of scores from 19.5 – 36, suggested the visuo-constructional skills of the participants were within age expectations.

Performance on the delayed procedure revealed the mean score was 14.8, *SD* = .92, with a range of 4.5 – 30. The majority of participants performed within two standard deviations of the mean and only one participant obtained a score below age expectations.

Table 3.09: The mean, standard deviation, and the range of performances on the Rey Complex Figure Test (n = 44).

Rey Complex Figure Test	*M*	*SD*	Range
Copy Score	33.8	.48	19.5 – 36
Delayed Recall	14.8	.92	4.5 – 30

3.3.3.2. The Visual Reproduction subtest

As this study aimed to develop a scoring system for an alternative form of the Visual Reproduction subtest of the Wechsler Memory Scale – Revised that was based on the scoring principles of Clark (2000) research, the Visual Reproduction subtest of the Wechsler Memory Scale – Revised was scored using the Original Scoring System published in the Wechsler Memory Scale – Revised manual (Wechsler, 1987) and the Revised Scoring System developed by Clark (2000).

3.3.3.2.1. The Original Scoring System

Table 3.10 shows that on the Original Scoring System the immediate mean score on Card A was 5.98, *SD* = .88, and comparable with the mean of 5.50, *SD* = 1.42 on Card B. Even though the range of scores was greater for Card B than for Card A, the inter-quartile ranges were similar for both. The mean score for Card C was 7.77, *SD* = .89. The range of scores was 6 – 9 scores but only one point differentiated the upper and lower inter-quartile range. The immediate mean score for Card D was 15.42, *SD* = 2.06 with the inter-quartile score range gathered towards the upper end of the score range of 10 – 18 points.

The mean Total Immediate Recall score on the Original Scoring System was 34.71, *SD* = 3.09. The range of scores obtained were 28 – 40, falling between the 18th and 98th percentile. The inter-quartile scores were grouped within the upper end of the score range, with ceiling effects evident. This indicated that the participants performed well, but no perfect scores were obtained when the Original Scoring System was applied.

Table 3.10: Immediate recall performances on the Original Scoring System of the Visual Reproduction subtest of the Wechsler Memory Scale – Revised (n = 44).

Original Scoring System	*M*	*SD*	Range	Inter-quartile Range		
				25	50	75
<u>Immediate Recall</u>						
Card A	5.98	.88	4 – 7	5	6	7
Card B	5.50	1.42	1 – 7	5	6	7
Card C	7.77	.89	6 – 9	7	8	8
Card D	15.42	2.06	10 – 18	14	16	17
Total Score	34.71	3.09	28 – 40	33	35	37

On the Original Scoring System, Table 3.11 shows the mean delayed recall score on Card A of 5.39, *SD* = 1.83, was comparable with the mean of 5.32, *SD* = 1.46, on Card B. The range of scores was similar for both Card A and Card B with the inter-quartile range also comparable. The mean score for Card C was 6.89, *SD* = 1.90 and the range of scores was 6 – 9, with only two points between the upper and lower inter-quartile range. The delayed mean score for Card D was 14.14, *SD* = 2.98 with inter-quartile range gathered towards the upper end of the range of 0 – 18 scores.

The mean Total Delayed Recall score on the Original Scoring System was 31.82, *SD* = 4.79, with a range of 16 – 39 points falling between the 7th and 96th percentile. The inter-quartile range was also grouped within the upper end of the score range, but ceiling effects were not evident. This indicated that the participants in this study performed well on the designs after a time delay, but no perfect scores were obtained when the Original Scoring System was applied.

Table 3.11: Delayed recall performances on the Original Scoring System of the Visual Reproduction subtest of the Wechsler Memory Scale – Revised (n = 44).

Original Scoring System	*M*	*SD*	Range	Inter-quartile Range		
				25	50	75
Delayed Recall						
Card A	5.39	1.83	0 – 7	5	6	6
Card B	5.32	1.46	1 – 7	5	6	6
Card C	6.89	1.90	0 – 9	6	7	8
Card D	14.14	2.98	0 – 18	13	14	16
Total Score	31.82	4.79	16 – 39	29	32	35

3.3.3.2.2. The Revised Scoring System

Table 3.12 shows that on the Revised Scoring System the mean immediate recall score on Card A of 18.7, *SD* = .67, was the highest, the mean score of 16.3, *SD* = 1.58, on Card B was the lowest, and the mean score for Card C of 17.8, *SD* = 1.61, was comparable to the mean score of 17.8, *SD* = 1.66, on Card D. The range of scores 17 – 20 on Card A was smaller than on the other three designs, but the score range of 13 – 20 for Card B, Card C and Card D was comparable. The inter-quartile ranges approached ceiling levels at the 25th percentile for Card A, Card C and Card D and although the inter-quartile range was lower for Card B, the score difference between the 25th and 75th percentile for all four designs was a maximum of two points.

The mean Total Immediate Recall score on the Revised Scoring System was 70.7, *SD* = 3.17. The range of scores was 62 – 77 and at the upper end of the score range, but no ceiling effects were evident. This indicated that the participants in this study of Low Average or above intelligence performed well after an initial presentation, but no perfect scores were obtained when the Revised Scoring System was applied.

Table 3.12: Immediate recall performances on the Revised Scoring System of the Visual Reproduction subtest of the Wechsler Memory Scale – Revised (n = 44).

Revised Scoring System	*M*	*SD*	Range	Inter-quartile Range		
				25	50	75
<u>Immediate Recall</u>						
Card A	18.7	.67	17 – 20	18	19	19
Card B	16.3	1.58	13 – 20	15	16	17
Card C	17.8	1.61	13 – 20	17	18	19
Card D	17.8	1.66	13 – 20	17	18	19
Total Score	70.7	3.17	62 – 77	69	71	73

Table 3.13 shows that on the Revised Scoring System the mean delayed recall score for Card A of 16.8, *SD* = 5.43, was similar to the mean score of 16.0, *SD* = 1.68 on Card B, the mean score of 16.1, *SD* = 3.81 on Card C, and the mean score of 16.8, *SD* = 3.20 on Card D. The score range for Card B was smaller than on the other designs at 12 – 20 with ceiling effects demonstrated. The score range of 0 – 20 for Card B, Card C and Card D was comparable but floor and ceiling effects were evident. Although the inter-quartile scores were lower for Card B, the score difference was small between the 25th and 75th percentile for all four designs with a maximum of two points.

Table 3.13: Delayed recall performances on the Revised Scoring System of the Visual Reproduction subtest of the Wechsler Memory Scale – Revised (n = 44).

Revised Scoring System	*M*	*SD*	Range	Inter-quartile Range		
				25	50	75
<u>Delayed Recall</u>						
Card A	16.8	5.43	0 – 20	18	18	19
Card B	16.0	1.68	12 – 20	15	16	17
Card C	16.1	3.81	0 – 20	15	17	19
Card D	16.8	3.20	0 – 20	16	17	19
Total Score	65.6	8.18	32 – 76	64	68	70

On the Revised Scoring System, the mean Total Delayed Recall score was 65.6, *SD* = 8.18. The range of scores was 32 – 76 with performances grouped towards the upper end of the range, but no ceiling effects were evident. This indicated that the participants in this study performed well after a time delay, but no perfect scores were obtained when the Revised Scoring System was applied.

3.3.4. Speed of Information Processing Tests

The Speed and Comprehension of Language Processing test and the Symbol Search subtest were used to determine the stability of performances between Session One and Session Two in a repeated measures design. It was expected that participants would perform better on session two due to practice effects, so a minimum processing speed score was established in the first session that set the benchmark in order for session two to proceed. This benchmark was set to ensure that extraneous factors such as anxiety, stress and fatigue were not influencing test performance, which was met by all participants.

Descriptive analyses and paired sample *t*-tests were utilised to compare the raw scores on the Speed and Comprehension of Language Processing test and the Symbol Search subtest between Session One and Session Two. Histograms were used to describe the distribution of score gains of the sample and scatter-plots examined the relationship between score gains and overall level of intellectual ability.

3.3.4.1. The Speed and Comprehension of Language Processing test

Participants raw scores on the Speed and Comprehension of Language Processing test were adjusted for age and summed to reveal a mean score of 62.1, *SD* = 2.81, on Session One and a higher mean score of 77.0, *SD* = 3.08 on Session Two. The mean score increase of 14.9, *SD* = 9.72, was statistically significant, $t(43) = 10.16$ $p = .0005$. The range of scores was relatively equivalent between Session One, 14 – 100, and Session Two, 16 – 100 but a large range of 0 – 41 points was gained between the two sessions, as shown in Table 3.14.

Table 3.14: The Speed and Comprehension of Language Processing test performances on Session One and Session Two and score gains on repeat administration (n = 44).

SCOLP	Session One			Session Two			Score Gain		
	M	*SD*	Range	*M*	*SD*	Range	*M*	*SD*	Range
Raw Score	62.1	2.81	14 – 100	77.0	3.08	16 – 100	14.9	9.72	0 – 41

Figure 3.04 highlights the distribution in score gains between Session One and Session Two on the Speed and Comprehension of Language Processing test.

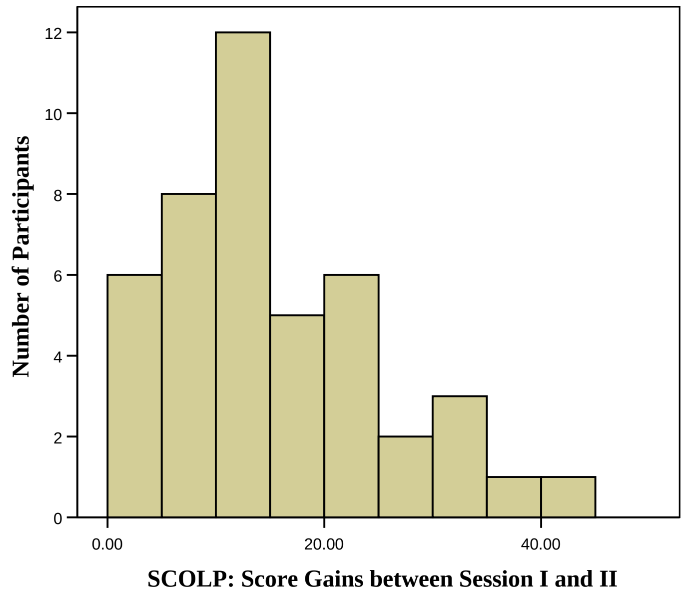


Figure 3.04: The distribution of score gains between Session One and Session Two on the Speed and Comprehension of Language Processing test.

Figure 3.04 illustrates that although a large distribution in score gains from 0 – 41 were obtained by participants between Session One and Session Two on the Speed and Comprehension of Language Processing test. The majority of participants gained 12 – 14 points, with only one participant obtaining the same score, and only a few gaining more than 25 points on the second session.

3.3.4.2. The Symbol Search subtest

On the Symbol Search subtest, the results revealed a mean raw score of 35.1, SD = 1.32, on Session One was a higher mean raw score of 39.6, SD = 1.32 on Session Two. The mean gain in raw scores of 5.07, SD = 3.84, was statistically significant, $t(43) = 9.55$, $p = .0005$. Although the range of raw scores was slightly lower in Session One, 18 – 56, than in Session Two 23 – 67, the magnitude of scores between the sessions was equivalent. A wide range of 0 – 17 points was gained between the two sessions, as shown in Table 3.15.

The raw scores of the Symbol Search subtest were corrected for age according to the normative data in the Wechsler Adult Intelligence Scale – Third Edition manual which revealed a mean scaled score of 10.7, *SD* = 3.00 on Session One and a higher mean scaled score of 12.5, *SD* = 3.11, on Session Two. The mean scaled score gain of 1.77, *SD* = 1.22 was statistically significant, $t(43) = 9.66$, $p < .0005$.

A large range of scaled scores on Session One, 6 – 18, and Session Two, 7 – 9, were obtained, but the magnitude of the scores between the sessions was equivalent. A wide range of 0 – 5 scaled score were gained between the sessions, as shown in Table 3.11.

Table 3.15: The Symbol Search subtest raw scores on Session One and Session Two and score gain on repeat administration (n = 44).

Symbol Search	Session One			Session Two			Score Gain		
	M	*SD*	Range	*M*	*SD*	Range	*M*	*SD*	Range
Raw Score	35.1	1.32	18 – 56	39.6	1.32	25 – 67	5.07	3.84	0 – 17

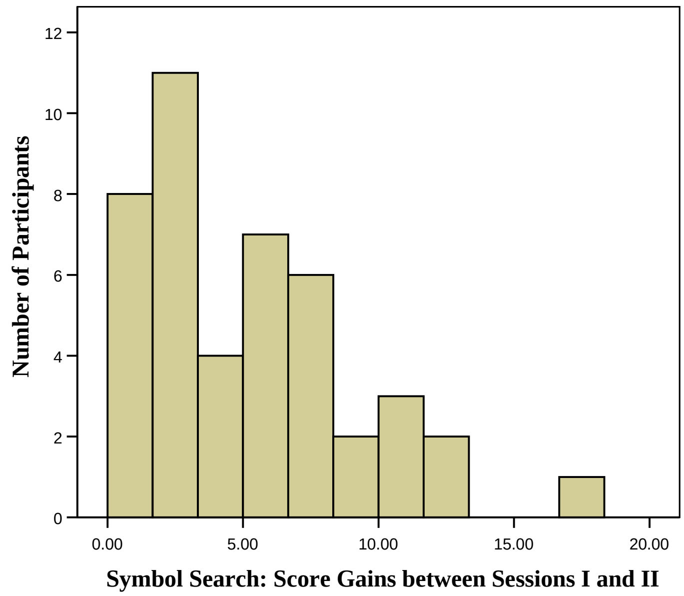


Figure 3.05: The distribution of raw score gains on the Symbol Search subtest between Session One and Session Two.

Figure 3.05 shows that a distribution in score gains ranged from 0 – 17 were obtained by participants between sessions. The majority of participants gained less than five points with only one participant with a comparable performance and one gaining an extra 17 points on the second session.

3.3.5. The Alternative Scoring System for the Alternative Form of the Visual Reproduction subtest of the Wechsler Memory Scale – Revised

The performances of each of the four designs of the alternative form developed in this study were analysed. The means, standard deviations and score distributions, skewness and kurtosis for the immediate recall performances on each of the four designs were conducted.

3.3.5.1. Immediate Recall Performances on the Alternative Scoring System

Table 3.16 shows on the Alternative Scoring System the mean immediate recall score of 16.6, *SD* = 3.27, for Design A was comparable to the mean performance of 16.3, *SD* = 2.95 on Design B, with a similar range of scores, 8 – 20, and 6 – 20 respectively. The mean performance of 17.9, *SD* = 1.74 was highest on Design C, with the smallest range of scores 12 – 20 obtained on this design. The lowest mean score of 13.4, *SD* = 4.10, was on Design D, which also had a broad range of scores 4 – 20, and a wider inter-quartile range than the other three designs. The inter-quartile range for Design C was small and approached ceiling level at the 25^{th} percentile. Although the inter-quartile range was lower for Design A and Design B, the score differences were small with a maximum of five points between the 25^{th} and 75^{th} percentile.

On the Alternative Scoring System the mean Total Immediate Recall score was 64.2, *SD* = 8.84. A relative broad range of scores 34 – 77 was obtained, but no floor or ceiling effects were evident. The inter-quartile range indicated a 14 point difference between the 25^{th} and 75^{th} percentile. This indicated that the participants in this study of Low Average or above intelligence performed reasonably well on the Alternative Scoring System after an initial presentation of the designs.

The immediate recall score distributions for Design A, Design B and Design C of the Alternative Form revealed a negatively skewed, with a more peaked and narrower distribution, which significantly departed from normality (see Table 3.16). However, the distribution of scores on Design D were slightly negatively skewed, skewness = -.21 and flatter than normal distribution, kurtosis = -.57, but did not depart significantly from normality.

Table 3.16: The means, standard deviations, and immediate recall score distributions of the Alternative Scoring System of the Alternative Form (n = 44).

Alternative Scoring System	*M*	*SD*	Range	Inter-quartile Range 25	50	75	Skewness	Kurtosis
<u>Immediate Recall</u>								
Design A	16.6	3.27	8 – 20	15	18	19	-1.16	.34
Design B	16.3	2.95	6 – 20	14	17	19	-1.30	2.47
Design C	17.9	1.74	12 – 20	17	18	19	-1.26	1.93
Design D	13.4	4.10	4 – 20	10	13	17	-.21	-.57
Total Score	64.2	8.84	34 – 77	58	64	72	-.89	1.71

Examination of the Total Immediate Recall score distribution revealed a mildly negatively skewed, skewness = -.89, but more sharply peaked and narrower distribution than the normal curve, kurtosis = 1.71, but this did not depart significantly from normality.

Figures 3.06 – 3.09 illustrate the immediate recall score distributions obtained on each of the Alternative Designs of the Alternative Form.

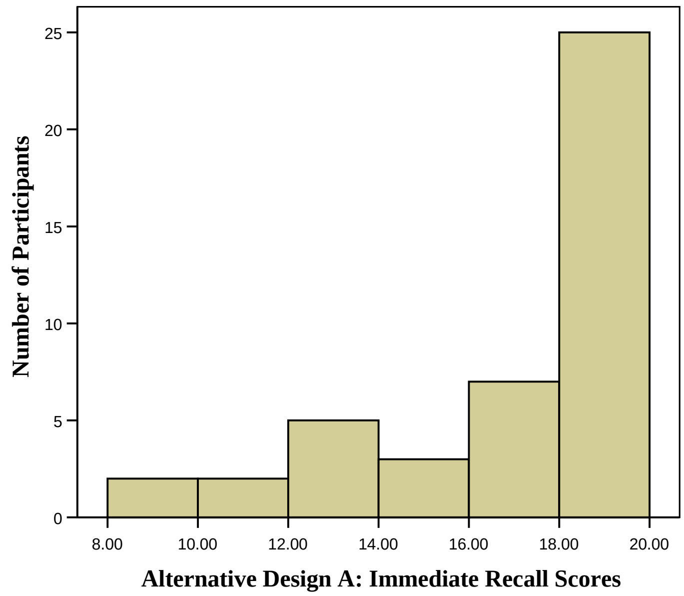


Figure 3.06: The immediate recall score distributions of Design A of the Alternative Form.

Figure 3.06 highlights the majority of participants scored 18 points or higher on Design A with ceiling effects evident. Only a few participants scored below 10 but no floor effects were demonstrated.

Figure 3.07: The immediate recall score distributions of Design B of the Alternative Form.

Figure 3.07 highlights the majority of participants scored 18 points or higher on Design B with ceiling effects evident. Only two participants scored below 10, but no floor effects were demonstrated.

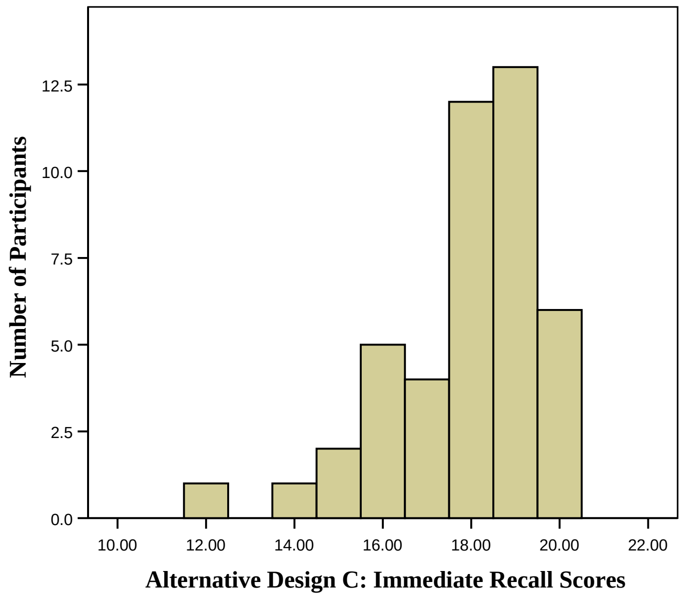


Figure 3.08: The immediate recall score distributions of Design C of the Alternative Form.

Figure 3.08 highlights the majority of participants scored 18 points or higher on Design C with ceiling effects evident, but no scores lower than 12 were obtained.

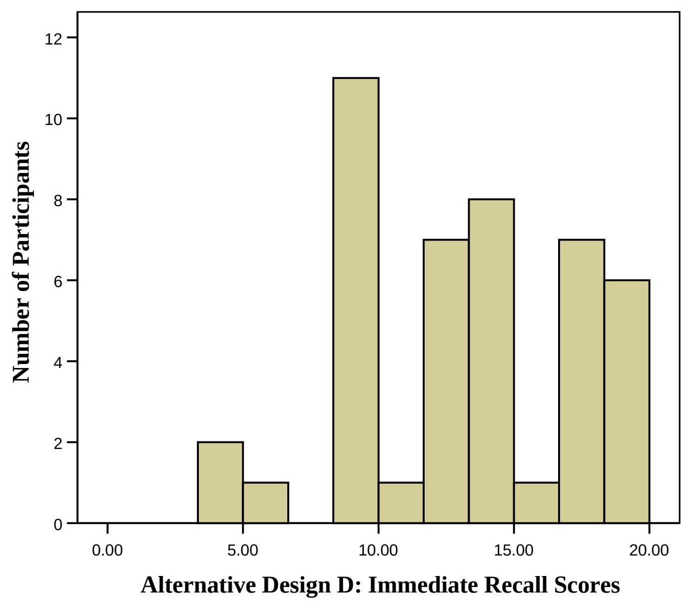


Figure 3.09: The immediate recall score distributions of Design D of the Alternative Form.

Figure 3.09 highlights the majority of participants scored 10 points or higher on Design D with ceiling effects evident. Only a few participants obtained a score lower than 10, but no floor effects were demonstrated.

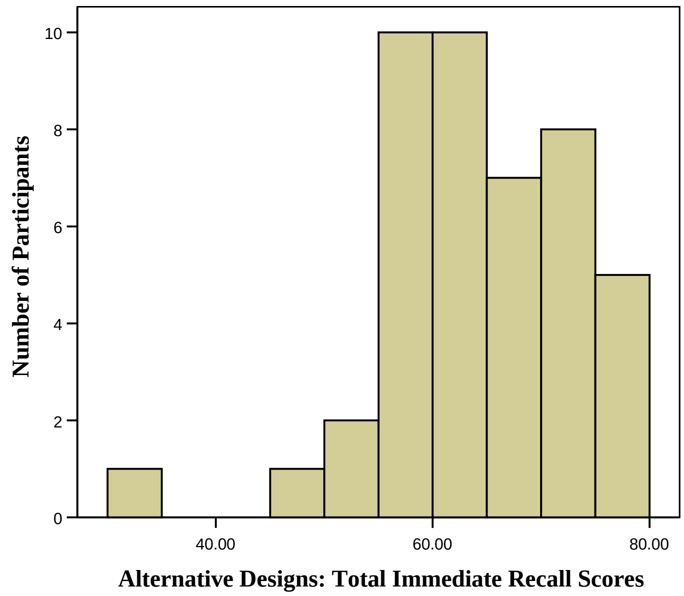


Figure 3.10: The Total Immediate Recall score distributions of the Alternative Designs of the Alternative Form

Figure 3.10 highlights the majority of participants performed well, scoring between 55 – 65 points or higher on the Total Immediate Recall score of the Alternative Scoring System. Only one participant performed below the midpoint of 40, but no floor or ceiling effects were evident.

3.3.5.1.1. Total Immediate Recall scores and Full Scale IQ scores

Figure 3.11 illustrates the relationship between the distribution of Total Immediate Recall scores of the Alternative Scoring System and general intellectual ability.

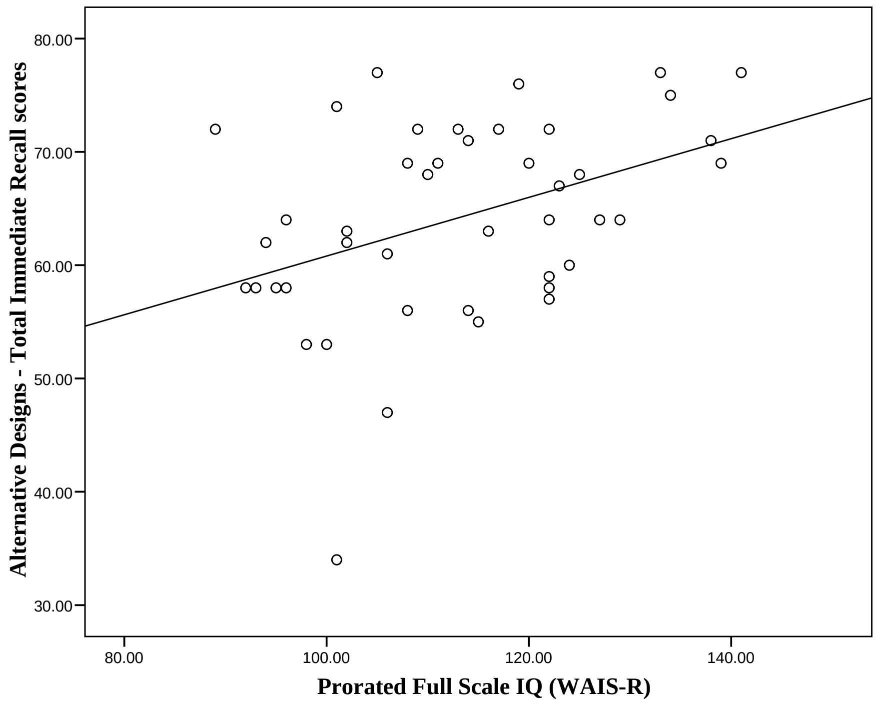


Figure 3.11: The distribution of the Total Immediate Recall scores of the Alternative Scoring System and general intellectual ability.

Figure 3.11 revealed participants with higher intellectual abilities performed better than those with lower levels of general functioning. Generally, individuals with Full Scale IQ within the Superior range and above obtained a score greater than 50 out of the maximum of 80 points. The lower scores were generally obtained by individuals with an IQ in the Average range, but performances at lower levels of ability were more widespread with some high scores also demonstrated. That is, a moderate relationship between overall intellectual ability and Total Immediate Recall performance on the Alternative Scoring System was evident.

3.3.5.2. Delayed Recall performances on the Alternative Scoring System

Table 3.17 shows on the Alternative Scoring System the mean delayed recall score of 9.00, *SD* = 7.46 for Design A was similar to the mean performance of 8.11, *SD* = 7.94 on Design B and the mean score of 10.3, *SD* = 6.20. The mean score of 17.4, *SD* = 2.46 was much higher on Design C than the other three designs. The range of scores for Design A, Design B and Design D were comparable with both floor and ceiling effects evident. The range of scores 7 – 20, on Design C reached ceiling levels but floor effects were not demonstrated. The inter-quartile range for Design A and Design B was broad but floor effects were evident at the 25th percentile. In contrast, the inter-quartile range for Design C was small, with ceiling effects approached at the 50th percentile. A broader inter-quartile range was evident on Design D than on Design C with more widespread performances between the 25th and 75th percentile but no ceiling or floor effects were evident.

Table 3.17: The means, standard deviations, and delayed recall score distributions for the Alternative Scoring System of the Alternative Form (n = 44).

Alternative Scoring System	*M*	*SD*	Range	Inter-quartile Range 25	50	75	Skewness	Kurtosis
<u>Delayed Recall</u>								
Design A	9.00	7.46	0 – 20	0	9	17	.08	-1.51
Design B	8.11	7.94	0 – 20	0	8	16	.15	-1.77
Design C	17.4	2.46	7 – 20	16	18	19	-2.02	6.37
Design D	10.3	6.20	0 – 20	5	10	15	-.27	-.98
Total Score	44.8	13.9	16 – 80	35	45	53	.31	.02

On the Alternative Scoring System, the mean Total Delayed Recall score was 44.8, *SD* = 13.9. A relative broad range of scores 16 – 80 was obtained but ceiling effects were evident. The inter-quartile range indicated an 18 point difference between the 25th and 75th percentile. This indicated that the participants in this study of Low Average or above intelligence performed reasonably well on the Alternative Scoring System following a time delay.

Analysis of the score distributions for the delayed recall of Design A, Design B and Design D of the alternative form revealed a slightly positively skewed, with a flatter than normal distribution, but this did not depart significantly from normality. On the other hand, Design C was negatively skewed, with a more peaked and narrow distribution than the normal curve which significantly departed from normality.

Examination of the scores on Total Delayed Recall of the Alternative Scoring System revealed a mildly positively skewed, skewness = .31, and a relatively normal distribution, kurtosis = .02, as shown in Table 3.17.

Figures 3.12 – 3.15 illustrate the distribution of scores obtained on each of the four Alternative Designs of the Alternative Form.

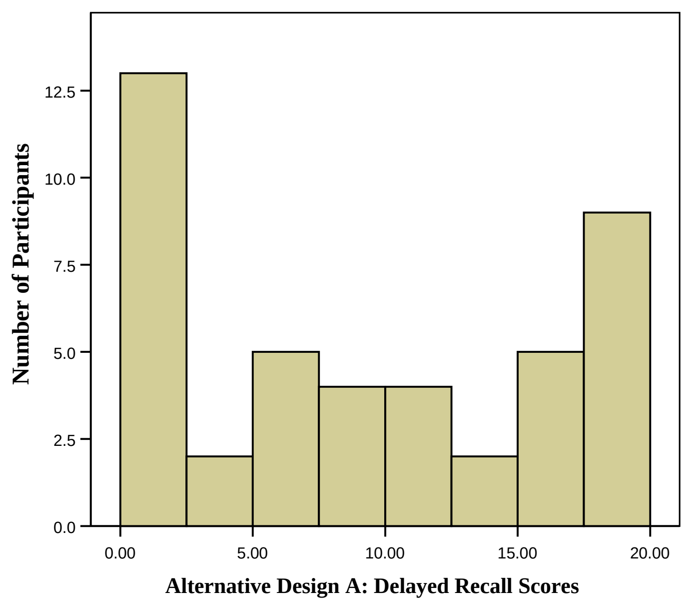


Figure 3.12: The delayed recall score distributions of Design A of the Alternative Form.

Figure 3.12 highlights the range of scores on delayed recall of Design A was widely distributed, with a tendency of scores to group towards the two extremes of the score range with a majority performing at lower end of the range.

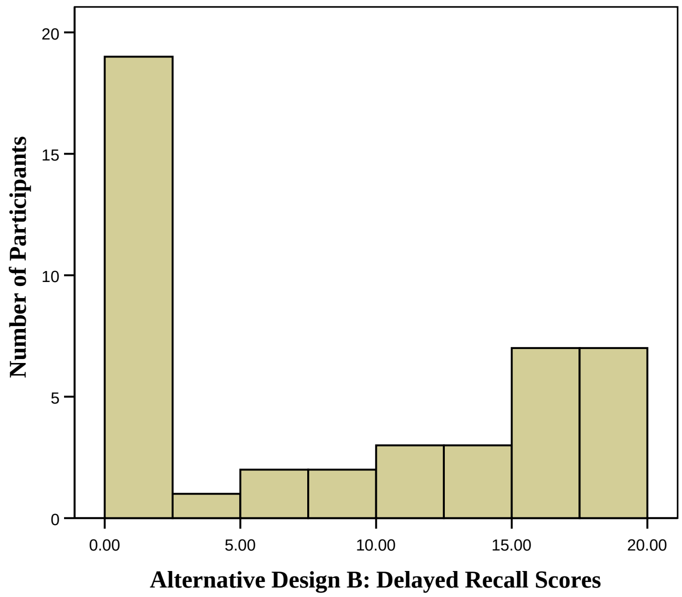


Figure 3.13: The delayed recall score distributions of Design B of the Alternative Form.

Figure 3.13 illustrates that the majority of participants obtained a score of zero on delayed recall of Design B. However, performances at both ends of the score range were evident.

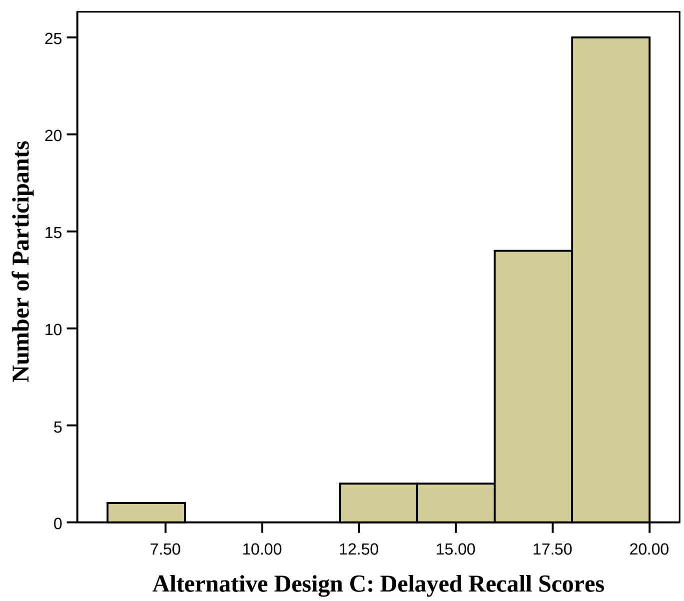


Figure 3.14: The delayed recall score distributions of Design C of the Alternative Form.

Figure 3.14 highlights that the majority of participants obtained a score greater than 15 on Design C approaching ceiling levels, but no floor effects were revealed.

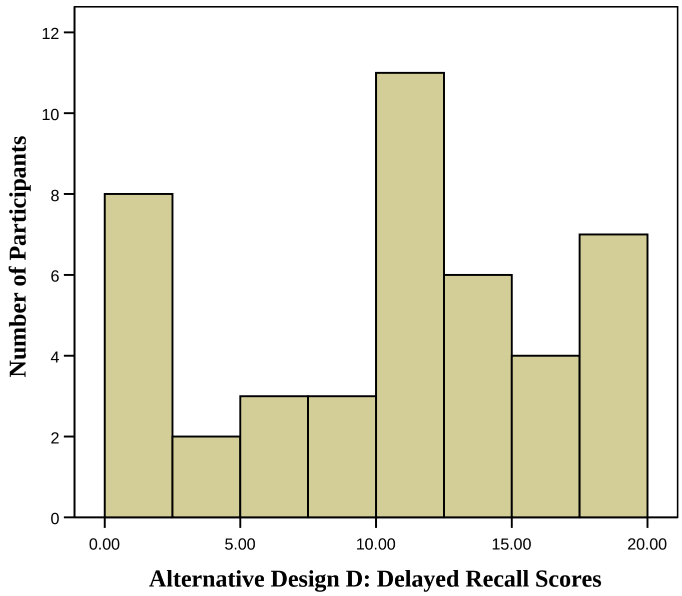


Figure 3.15: The delayed recall score distributions of Design D of the Alternative Form.

Figure 3.15 indicates that the majority of participants obtained a score of 10 – 13 on Design D. Approximately the same number of individuals performed at either score extremes with both floor and ceiling effects evident.

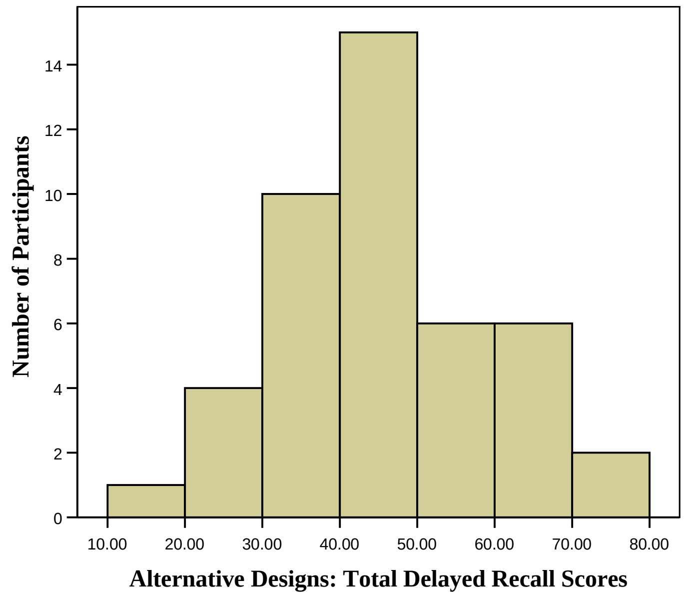


Figure 3.16: The Total Delayed Recall score distributions of the designs of the Alternative Form.

Figure 3.16 highlights the majority of participants scored between 40 – 50 points on the Total Delayed Recall score. A small number of participants performed at the two score extremes with only ceiling effects demonstrated.

3.3.5.1.2. Total Delayed Recall scores and Full Scale IQ scores

Figure 3.17 illustrates the distribution of Total Delayed Recall scores in comparison to the Full Scale IQ scores.

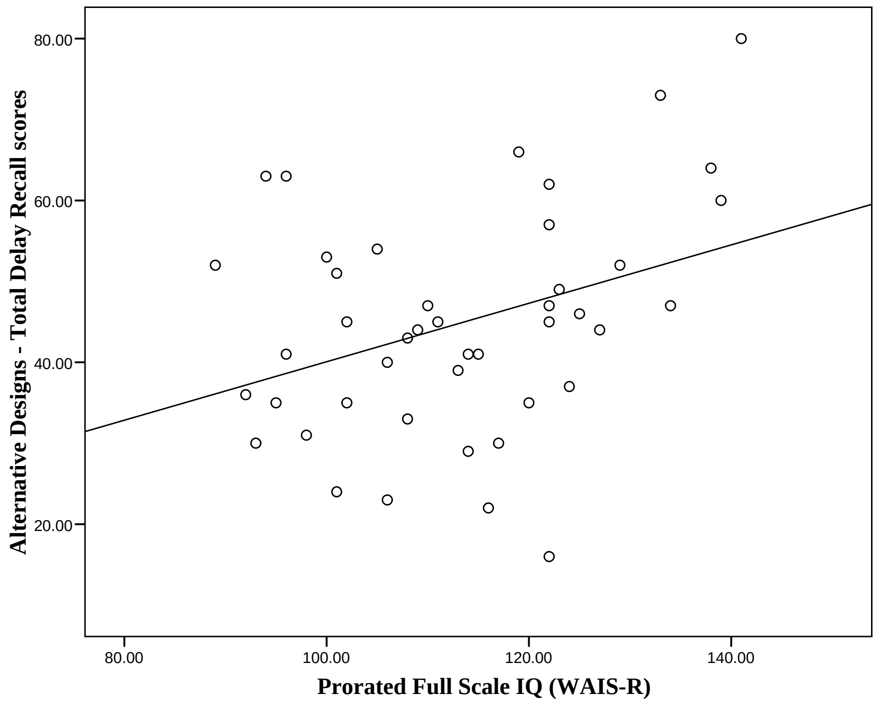


Figure 3.17: The distribution of total delayed scores of the Alternative Form and general intellectual ability.

Figure 3.17 highlights a tendency for participants of higher intellectual ability to perform better than those of lower levels of general functioning. That is, a weak relationship between overall intellectual ability and Total Delayed Recall performance on the Alternative Scoring System was evident.

3.4. Hypothesis Testing

3.4.1. Reliability of the Alternative Form

The potential reliability of the scoring system was measured by evaluating the internal consistency of the scoring items for each of the four designs. The internal consistency of the Alternative Scoring System for immediate and delayed recall for the Alternative Designs was evaluated by computing Cronbach's alpha.

Table 3.18: Internal consistency of the Alternative Scoring System for immediate and delayed recall scores of the Alternative Designs (n = 44).

Alternative Scoring System	Cronbach's Alpha	
	Immediate	Delayed
Design A	.87	.96
Design B	.80	.93
Design C	.61	.66
Design D	.83	.91
Total Score	.90	.94

Table 3.18 shows the internal consistency reliability coefficient for the Immediate Recall scores of Designs A, B and D was high, but only moderate reliability for Design C was evident. The internal consistency of the Total Immediate Recall score was also high which suggested good internal reliability of the Alternative Scoring System in this study.

Similarly, the delayed recall internal consistency coefficients for Designs A, B and D were high, but only moderate reliability for Design C was evident. The internal consistency of the Total Delayed Recall score was high and reflected favourably on the reliability of the Alternative Scoring System.

3.4.2. Relationship between the Alternative Form and the Visual Reproduction subtest of the Wechsler Memory Scale – Revised

Given the small sample size of this study, and the greater distribution of scores obtained on the combined total recall score of the four designs (i.e. from 20 to 80 points), only the Total Immediate Recall and Total Delayed Recall scores were utilised in making comparisons between the Alternative Scoring System and other test performances. As the distribution of Total Immediate Recall and Total Delayed Recall scores on the

Alternative Scoring System did not depart significantly from normality, the relationship for all comparisons were calculated using Pearson’s r correlation. However, a more conservative correlation using Spearman’s Rho was also conducted.

Table 3.19: Correlation between the Total Immediate Recall scores on the Alternative Scoring System and the Original Scoring System and Revised Scoring System (n = 44).

Total Immediate Recall	Alternative		Original		Revised	
	r	r_s	r	r_s	r	r_s
Alternative Scoring System	1	1				
Original Scoring System	.51**	.45**	1	1		
Revised Scoring System	.47**	.43**	.79**	.82**	1	1

** Correlation significant, p < .01 (2 - tailed).

Table 3.19 illustrates that a significant positive moderate correlation was found between the Total Immediate Recall score on the Alternative Scoring System with the Original Scoring System r = .51, p < .01; r_s = .45, p < .01, which significantly accounted for 25% of the variance in performance, $F(1,43) = 15.06$, p = .0005.

A significant positive moderate correlation was also found between the Total Immediate Recall score on the Alternative Scoring System and the Revised Scoring System, r = .47, p < .01; r_s = .43, p < .01, which significantly accounted for 20% of the variance in performance, $F(1,43) = 11.68$, p = .001.

A significant positive strong relationship between the Original Scoring System and the Revised Scoring System was found, r = .79, p < .01; r_s = .82, p < .01, which suggested the Immediate Recall scores on the two scoring systems were highly associated.

Table 3.20: Correlation between the Total Delayed Recall scores on the Alternative Scoring System and the Original Scoring System and Revised Scoring System (n = 44).

Total Delayed Recall	Alternative		Original		Revised	
	r	r_s	*r*	r_s	*r*	r_s
Alternative Scoring System	1	1				
Original Scoring System	.43**	.37*	1	1		
Revised Scoring System	.23 +	.17 +	.83**	.80**	1	1

** Correlation significant p < .01 (2 - tailed) + correlation not significant p > .05 (2 – tailed)

Table 3.20 illustrates a significant positive moderate relationship was found between the Total Delayed Recall score on the Alternative Scoring System and the Original Scoring System, $r = .43$, $p < .01$; $r_s = .37$, $p < .05$, which significantly accounted for 17% of the variance in performance, $F(1,43) = 9.54$, $p = .004$.

The Delayed Recall score on the Alternative Scoring System had a weak and positive association with the Revised Scoring System which was not significant, $r = .23$, $p > .05$; $r_s = .17$, $p > .05$.

The relationship between the Original Scoring System and the Revised Scoring System on Delayed Recall was significant and strong, $r = .83$, $p < .01$, which suggested the Delayed Recall scores on the two scoring systems were highly associated.

3.4.3. Validity

The validity of the Alternative Scoring System was evaluated to determine whether the measure adequately assessed the hypothetical construct it was developed to measure. The convergent and discriminant validity of the Alternative Scoring System was examined by determining its relationship with well established measures of verbal memory and learning (Logical Memory and Rey Auditory Verbal Learning Test), and another non-verbal memory measure (Rey Complex Figure Test).

As previously indicated, given the Total Immediate Recall score and the Total Delayed Recall scores on the Alternative Scoring System did not depart significantly from normality, the relationship between these measures was analysed used Pearson's r correlations.

3.4.3.1. Convergent Validity

In order to provide support for its validity, the delayed memory score of the Alternative Scoring System was expected to share a moderate – strong association with the Rey Complex Figure Test.

Table 3.21: The relationship between the Total Delayed Recall Score of the Alternative Scoring System and the Rey Complex Figure Test (n = 44).

	Rey Complex Figure Test Delayed Recall
Alternative Scoring System	
Total Delayed Recall	.40**

** Correlation significant p < .01 (2 - tailed).

Table 3.21 revealed a significant positive moderate correlation between Total Delayed Recall score of the Alternative Scoring System and the Delayed Recall score on the Rey Complex Figure Test, r = .40, p < .01, which significantly accounted for 14% of the variance in performance, F(1,43) = 7.76, p = .008.

Table 3.22: The association between the Alternative Scoring System and the Rey Complex Figure Test (n = 44).

	Rey Complex Figure Test Copy Score
Alternative Scoring System	
Total Immediate Recall	.61**

** Correlation significant p < .01 (2 - tailed).

Table 3.22 revealed a significant moderate positive association with the Total Immediate Recall score of the Alternative Scoring System and the Copy score of the Rey Complex Figure Test, $r = .61$, $p < .01$, which significantly accounted for 36% of the variance in performance, $F(1,43) = 26.62$, $p = .0005$.

3.4.3.2. Discriminant Validity

In order to provide support for its discriminant validity, the Total Delayed Recall score of the Alternative Scoring System was expected to share a weak relationship with the Logical Memory subtest and the Rey Auditory Verbal Learning Test.

Table 3.23: The association between the Total Delayed Recall score on the Alternative Scoring System and the Logical Memory and Rey Auditory Verbal Learning Test delayed recall scores (n = 44).

	Logical Memory Delayed Recall	RAVLT Delayed Recall
Alternative Scoring System		
Total Delayed Recall	.27	.07

No significant correlations at p > .05 (2 – tailed).

Table 3.23 demonstrates a weak relationship between the Total Delayed Recall score of the Alternative Scoring System and the Delayed Recall scores on the Logical Memory subtest, *r* = .27, *p* > .05, and the Rey Auditory Verbal Learning Test, *r* = .07, *p* > .05, which was not significant.

Table 3.24: The association between the Total Immediate Recall score of the Alternative Scoring System and the Immediate Recall score of the Logical Memory subtest and the Learning score of the Rey Auditory Verbal Learning Test (n = 44).

	Logical Memory Immediate Score	RAVLT Learning Score
Alternative Scoring System		
Total Immediate Recall	.19	.15

No significant correlations at p > .05 (2 – tailed).

Table 3.24 illustrates the relationship between the Total Immediate Recall score of the Alternative Scoring System and the Immediate Recall score on the Logical Memory subtest and the Learning score on the Rey Auditory Verbal Learning Test were not significant and the correlations were weak, *r* = .19, *p* > .05, and *r* = .15, *p* > .05 respectively.

3.5. Factors potentially influencing performance

3.5.1. Order of Test Administration

The order of administration of the alternative form was alternated between the two testing sessions for all participants. In examining whether test order effects on Total Immediate Recall and Total Delayed Recall scores of the Alternative Scoring System were evident, descriptive and independent samples *t*-test analyses were conducted.

Table 3.25: The Total Immediate Recall and Total Delayed Recall scores of the Alternative Scoring System and order of test administration (n = 44).

Session	%	Total Immediate Recall		Total Delayed Recall	
		M	*SD*	*M*	*SD*
Session One	48%	64.3	1.60	43.8	1.64
Session Two	52%	64.1	6.65	45.9	1.80
Score Difference		.17	2.70	2.07	4.24

Table 3.25 shows the mean Total Immediate Recall score when the Alternative Designs were administered in the first session of 64.3, *SD* = 1.60, was comparable to the mean score of 64.1, *SD* = 6.65 in the second session, with a mean difference of .17, *SD* = 2.70. The results indicated that the order of test administration did not significantly influence the Total Immediate Recall performance on the Alternative Scoring System, $t(42) = .06$, $p = .95$.

The mean Total Delayed Recall score when the Alternative Designs were administered in the first session of 43.8, *SD* = 1.64, was slightly lower than the mean score of 45.9, *SD* = 1.80, when administered in the second session, with a mean difference of 2.07, *SD* = 4.24. The results indicated that the order of test administration did not significantly influence the Total Delayed Recall performance on the Alternative Scoring System, $t(42) = .49$, $p = .63$.

3.5.2. Age

In order to determine if the Total Immediate Recall and Total Delayed Recall scores of the Alternative Scoring System varied as a function of age, descriptive analyses and a one-way Analysis of Variance (ANOVA) between the three age groups 25 – 34, 35 – 44 and 45 – 54 of this study were conducted. The age groups were devised to replicate the age bands published in the Wechsler Memory Scale – Revised manual.

Table 3.26: The Total Immediate Recall and Total Delayed Recall scores of the three age groups on the Alternative Scoring System (n = 44).

Alternative Scoring System	n	*M*	*SD*	Range
Total Immediate Recall				
25 – 34 age group	16	66.8	7.52	53 – 77
35 – 44 age group	17	64.5	7.44	53 – 77
45 – 54 age group	11	59.8	11.4	34 – 71
Total Delayed Recall				
25 – 34 age group	16	49.9	12.6	33 – 80
35 – 44 age group	17	41.9	14.7	16 – 73
45 – 54 age group	11	41.7	13.3	23 – 64

Table 3.26 indicates the mean Total Immediate Recall score decreased from 66.8, *SD* = 7.52 for the 25-34 age group, to a mean score of 64.5, *SD* = 7.44 for the 35 – 44 year olds, and a mean score of 59.8, *SD* = 11.4, for the 45 – 54 age group. Although the Total Immediate Recall score decreased as a function of age, age did not have a significant influence on performance for this group, $F(2,43) = 2.18$, $p = .23$.

Table 3.26 shows that the mean Total Delayed Recall score of 49.9, *SD* = 12.6, was higher for younger participants, the mean score of 41.9, *SD* = 14.7, for the 35 – 44 year olds, and the mean score of 41.7, *SD* = 13.3, for the 45 – 54 age group was comparable. Although younger participants performed better, age did not have a significant influence on the Total Delayed Recall score, $F(2,43) = 1.80$, $p = .18$.

3.5.3. Education

In order to determine if years of education influenced the Total Immediate Recall and Total Delayed Recall scores on the Alternative Scoring System, descriptive and a one-way ANOVA analyses was conducted.

Table 3.27: The Total Immediate Recall and Total Delayed Recall performances for the three levels of education on the Alternative Scoring System (n = 44).

Alternative Scoring System	n	*M*	*SD*	Range
<u>Total Immediate Recall</u>				
Below 12 years	16	61.31	10.95	34 – 77
12 years	15	65.87	8.16	53 – 77
Over 12 years	13	65.77	5.92	58 – 77
<u>Total Delayed Recall</u>				
Below 12 years	16	40.50	13.82	23 – 73
12 years	15	47.13	12.38	22 – 66
Over 12 years	13	47.30	15.42	16 – 80

Table 3.27 shows the mean Total Immediate Recall score of the Alternative Scoring System of 61.31, *SD* = 10.95 was lower for participants with less than 12 years of education than the mean score of 65.87, *SD* = 8.16, for participants with 12 years and the mean of 65.77, *SD* = 5.92 for over 12 years of education. Although the Total Immediate Recall score was lower for participants with less than 12 years of education and comparable with 12 years or higher, years of education did not significantly influence test performance for this group, $F(2,43) = 1.35$, $p = .27$.

The mean Total Delayed Recall score on the Alternative Scoring System of 40.50, *SD* = 13.82 was lower for participants with less than 12 years of education than the mean score of 47.13, *SD* = 12.38, for participants with 12 years, and the mean of 47.30, *SD* = 15.42 for over 12 years of education. Although the Total Immediate Recall score was lower for participants with less than 12 years of education and comparable with 12 years or higher, years of education did not significantly influence test performance for this group, $F(2,43) = 1.20$, $p = .31$.

3.5.4. Gender

In examining whether gender differences were evident on the Total Immediate Recall score and Total Delayed Recall score on the Alternative Scoring System and other measures, descriptive and independent sample *t*-test analyses were conducted.

Table 3.28: Performances of males (n = 24) and females (n = 20) on the Alternative Scoring System.

Alternative	Males		Females		Difference	
Scoring System	*M*	*SD*	*M*	*SD*	*M*	*SD*
Total Immediate Recall	67.88	6.33	59.75	9.50	8.13	2.40
Total Delayed Recall	47.92	14.33	41.00	12.73	6.92	4.13

Table 3.28 revealed a mean Total Immediate Recall score for males of 67.88, *SD* = 6.33, was greater than the mean score for females of 59.75, *SD* = 9.50. The mean difference of 8.13, *SD* = 2.40 was statistically significant, $t(42) = 3.39$, $p = .002$, indicating that males performed significantly better on Total Immediate Recall on the Alternative Scoring System than females.

Similarly, the mean Total Delayed Recall score for males of 47.92, *SD* = 14.33, was greater than the mean Total Delayed Recall score for females of 41.00, *SD* = 12.73. However, the mean score difference of 6.92, *SD* = 4.13, was not statistically significant, $t(42) = 1.68$, $p = .10$. That is, although the males generally performed better than females, a significant difference in performance between males and females was not evident on delayed recall.

3.5.5. Speed of Information Processing

Pearson r correlation analyses were conducted in order to determine the relationship between processing speed and Total Immediate Recall and Total Delayed Recall performances on the Alternative Scoring System.

The results revealed that the Total Immediate Recall score on the Alternative Form had a negative weak correlation with the Speed and Comprehension of Language Processing test which was not significant, $r = -.00$, $p > .05$.

Similarly the Total Delayed Recall score had a negative very weak correlation with the Speed and Comprehension of Language Processing test which was not significant, $r = -.06$, $p > .05$. This indicated that as performance on delayed recall increased the less of an association with the verbal processing task was evident.

Together, these results suggest that with the Total Immediate Recall and the Total Delayed Recall performances on the Alternative Scoring System shared very little commonality with verbal information processing ability.

Table 3.29: Pearson's r correlation coefficients between the Alternative Scoring System and the Symbol Search subtest (n = 44).

Measure	Symbol Search
Alternative Scoring System	
Total Immediate Recall	.32*
Total Delayed Recall	.16

* Correlation significant at p < .05 (2 - tailed).

Table 3.29 illustrates that the Total Immediate Recall score on the Alternative Scoring System had a weak but significant positive moderate correlation with the Symbol Search score, $r = .32$, $p < .05$. Although significant, the Total Immediate Recall score on the Alternative Scoring System only accounted for 8% of the variance in the Symbol Search subtest performance, $t(43) = 2.19$, $p = .15$.

A positive weak correlation between the Total Delayed Recall score on the Alternative Scoring System and the Symbol Search score was found but this was not significant, $r = .16$ $p > .05$.

Taken together, the results of correlation analyses suggested that the Alternative Scoring System shared commonality with the rate of non-verbal information processing on immediate recall but this relationship was not found on delayed recall.

3.5.6. Level of Intellectual Functioning

3.5.6.1. Alternative Scoring System and IQ scores

As general intellectual ability has a well known relationship with memory performance, the association between the Total Immediate Recall and Total Delayed Recall scores of the Alternative Scoring System and the Full Scale IQ, Verbal IQ and Performance IQ scores was investigated. Descriptive and Pearson's r correlation analyses were conducted. Multiple regression analyses were also conducted to determine how much of the variance in the Total Immediate Recall and Total Delayed Recall scores of the Alternative Scoring System was predicted by the Full Scale IQ, Verbal IQ, Performance IQ and individual subtest scores.

Table 3.30: The Pearson's r correlation coefficients of the Full Scale IQ, Verbal IQ and Performance IQ scores with the Total Immediate Recall and Total Delayed Recall performances on the Alternative Scoring System (n = 44).

WAIS – R	Alternative Scoring System	
	Total Immediate	Total Delayed
Full Scale IQ	.40**	.36*
Verbal IQ	.33*	.25+
Performance IQ	.39**	.40**

* Correlation significant p < .05 (2 - tailed). + No significant correlation p > .05.
** Correlation significant p < .01 (2 - tailed).

The results in Table 3.30 show that performance on Total Immediate Recall on the Alternative Scoring System shared a significant positive moderate relationship with overall intellectual functioning, r = .40, p < .01 significantly accounting for 14% in performance, $F(1,43) = 8.12$ $p = .007$.

Total Immediate Recall on the Alternative Scoring System demonstrated a significant positive weak association with both general verbal skills, r = .33, p < .05, and non-verbal ability, r = .39, p < .01, significantly accounting for 9% of the variance in the score, $F(1,43) = 5.40$ $p = .03$, and 13% of the variance in performances, $F(1,43) = 7.60$ $p = .009$, respectively.

Total Delayed Recall performance, indicated that the Alternative Scoring System shared a significant positive moderate association with overall intellectual ability, r = .36, p < .05, which significantly accounted for 11% of the variance in performance, $F(1,43) = 6.11$ $p = .02$. Similarly, the Alternative Scoring System had a significant positive moderate correlation with the Performance IQ score, r = .40, p < .01, significantly accounting for 14% of the variance in performance. In contrast, the

Alternative Scoring System had a positive weak correlation with Verbal IQ which was not significant, $r = .25$, $p > .05$.

The moderate relationships shared between Total Immediate Recall and Total Delayed Recall performances with measures of general intellectual ability were expected, as correlations between memory and IQ are well established. However, the relationships were moderate at best, which suggested that the Alternative Scoring System was not loading greatly on IQ skills. This lends support for the Alternative Form as a measure of memory rather than general intellectual ability in this study.

3.5.6.2. The Alternative Scoring System and IQ subtests

The association of the individual Wechsler Adult Intelligence Scale – Revised subtests with the Alternative Scoring System are shown below in Table 3.31.

Table 3.31: The Pearson's r correlation coefficients of the individual Wechsler Adult Intelligence Scale – Revised subtests with the Total Immediate Recall and Total Delayed Recall scores on the Alternative Scoring System (n = 44).

	Alternative Scoring System	
Subtest	Immediate	Delayed
Information	.25	.25
Vocabulary	.13	.19
Similarities	.44**	.22
Digit Span	.22	.23
Arithmetic	.37*	.14
Picture Completion	.18	.19
Object Assembly	.42**	.51**
Block Design	.63**	.52**
Coding	.16	.12

* Correlation significant p < .05 (2 - tailed)
** Correlation significant p < .01 (2 - tailed).

The analysis of the relationship between individual subtests and Total Immediate Recall performance on the Alternative Scoring System revealed a significant positive strong relationship with the Block Design subtest, $r = .62$, $p < .01$, a moderate correlation with the Object Assembly subtest, $r = .42$, $p < .01$ and the Similarities subtest, $r = .44$, $p < .01$. A significant positive weak correlation was found between the Alternative Scoring System and the Arithmetic subtest, $r = .37$, $p < .05$. The positive weak correlations between the Total Immediate Recall score on the Alternative Scoring System and the Information subtest, $r = .25$, $p > .05$, the Vocabulary subtest, $r = .13$, $p > .05$, the Digit Span subtest, $r = .22$, $p > .05$, the Picture Completion subtest, $r = .18$, $p > .05$ and the Coding subtest, $r = .16$, $p > .05$, were not significant.

Positive weak correlations between the Information subtest, $r = .25$, $p > .05$, Arithmetic subtest, $r = .14$, $p > .05$, Similarities, $r = .22$, $p > .05$, Digit Span, $r = .23$, $p > .05$, Vocabulary, $r = .19$, $p > .05$, Picture Completion subtest, $r = .19$, $p > .05$, and the Coding subtest, $r = .12$, $p > .05$, were evident on delayed recall but these relationships were not significant. Only two significant positive moderate relationships were revealed between the Total Delayed Recall performance on the Alternative Scoring System and the Block Design subtest, $r = .52$, $p < .01$, and the Object Assembly subtest, $r = .51$, $p < .01$.

3.5.6.3. Original Scoring System and IQ scores

As the Original Scoring System of the Visual Reproduction subtest is an established memory measure of memory function, its relationship with general intellectual ability was also examined. The was done in order to determine whether a similar pattern of relationships were shared between the Alternative Scoring System and the Original Scoring System was established.

Table 3.32: The relationship between the immediate and delayed performance on the Original Scoring System and Full Scale IQ, Verbal IQ, and Performance IQ scores (n = 44).

Measure	Original Scoring System Immediate	Original Scoring System Delayed
Full Scale IQ	.18	.22
Verbal IQ	.12	.05
Performance IQ	.20	.39**

** Correlation significant p < .01 (2 - tailed).

Table 3.32 revealed the Immediate Recall score of the Original Scoring System shared weak associations with overall intellectual ability, r = .18, p > .05, general verbal, r = .12, p > .05, and non-verbal skills r = .20, p > .05, which were not significant.

Similarly, Delayed Recall performance on the Original Scoring System demonstrated a positive weak relationship with overall intellectual functioning, r = .22, p > .05, and general verbal skills, r = .05, p > .05, which was not significant. However, a significant positive moderate association with Performance IQ, r = .39, p < .01, significantly accounting for 13% in performance, F(1,43) = 7.51, p = .009.

Original Scoring System and the Revised Scoring System of the Visual Reproduction subtest and IQ subtests

Analysis of the association of individual Wechsler Adult Intelligence Scale – Revised subtests and immediate and delayed performances of the Original Scoring System and the Revised Scoring System were also conducted and are shown in Table 3.33 and Table 3.34.

Table 3.33: The Pearson's r correlation coefficients of the individual Wechsler Adult Intelligence Scale – Revised subtests with the immediate recall performances on the Original Scoring System and Revised Scoring System (n = 44).

	Original Scoring System	Revised Scoring System
Subtest	Immediate	Immediate
Information	.13	.09
Vocabulary	.18	.10
Similarities	.18	.08
Digit Span	.10	.19
Arithmetic	.15	.21
Picture Completion	.22	.18
Object Assembly	.31*	.26
Block Design	.45**	.43**
Coding	.06	.00

* Correlation significant p < .05 (2 - tailed).
** Correlation significant p < .01 (2 - tailed).

On Immediate Recall, only two significant positive moderate relationships between the Original Scoring System and the subtests of the Wechsler Adult Intelligence Scale – Revised were indicated, namely with the Block Design subtest, $r = .45$, $p < .01$ and Object Assembly subtest, $r = .31$, $p < .05$. No significant relationships were revealed between Immediate Recall performances and the Information subtest, $r = .13$, $p > .05$, Vocabulary subtest, $r = .18$, $p > .05$, Similarities subtest, $r = .18$, $p > .05$, Digit Span subtest, $r = .10$, $p > .05$, Arithmetic subtest, $r = .15$, $p > .05$, Picture Completion subtest, $r = .22$, $p > .05$, and the Coding subtest, $r = .06$, $p > .05$.

The Revised Scoring System only demonstrated a significant positive moderate correlation with the Block Design subtest, r = .43, p < .01. A weak association was evident with the Information subtest, r = .09, p > .05, Vocabulary subtest, r = .10, p > .05, Similarities subtest, r = .08, p > .05, Digit Span subtest, r = .19, p > .05, Arithmetic subtest, r = .21, p > .05, Picture Completion subtest, r = .18, p > .05, Coding subtest, r = .09, p > .00, and the Object Assembly subtest r = .26, p > .05 but this was not significant.

Table 3.34: The Pearson's r correlation coefficients of the individual Wechsler Adult Intelligence Scale – Revised subtests with the delayed recall performances on the Original Scoring System and Revised Scoring System (n = 44).

Subtest	Original Scoring System Delayed	Revised Scoring System Delayed
Information	.12	.04
Vocabulary	-.12	-.24
Similarities	-.10	-.23
Digit Span	.10	-.03
Arithmetic	.27	.18
Picture Completion	.06	-.03
Object Assembly	.52**	.39**
Block Design	.50**	.35*
Coding	.15	.13

* Correlation significant p < .05 (2 - tailed).
** Correlation significant p < .01 (2 - tailed).

Only two significant relationships were revealed on Delayed Recall for the Original Scoring System and the Revised Scoring System, namely the Block Design subtest, r = .52, p < .01, and the Object Assembly subtest, r = .50, p < .01, respectively.

Four positive weak relationships were found between the Original Scoring System and the Digit Span subtest, $r = .10$, $p > .05$, Arithmetic subtest, $r = .27$, $p > .05$, Picture Completion subtest, $r = .06$, $p > .05$, and the Coding subtest, $r = .15$, $p > .05$. Similarly, positive weak relationships were evident between the Revised Scoring System and the Information subtest, $r = .04$, $p > .05$, Arithmetic subtest, $r = .18$, $p > .05$ and Coding subtest, $r = .13$, $p > .05$.

As scores on the Original Scoring System increased an inverse relationship was found with the Vocabulary subtest, $r = -.12$, $p > .05$, and Similarities subtest, $r = -.10$, $p > .05$. Concordantly, negative correlations between the Revised Scoring System and the Vocabulary subtest, $r = -.24$, $p > .05$, Similarities subtest, $r = -.23$, $p > .05$, Digit Span subtest $r = -.03$, $p > .05$, and Picture Completion subtest, $r = -.03$, $p > .05$.

3.6. Predictors of the Alternative Scoring System

In order to determine the best predictive model of performance on the Total Immediate Recall and Total Delayed Recall scores on the Alternative Scoring System, multiple regression analyses were conducted using the Enter method. Due to the small sample size of the study only variables that demonstrated a significant correlation with the Alternative Scoring System were included so that the power size needed in order to achieve significance would not be inflated.

3.6.1. Predictors of the Total Immediate Recall score on the Alternative Scoring System

Multiple linear regression analyses were conducted using the Enter method to determine which independent variables significantly predicted the Total Immediate Recall and Total Delayed Recall scores on the Alternative Scoring System.

In the regression analyses of the Alternative Scoring System, the independent variables were the Similarities subtest, Block Design subtest, Arithmetic subtest, Object Assembly subtest, Full Scale IQ, Verbal IQ, Performance IQ, Symbol Search subtest, and the Copy Score of the Rey Complex Figure Test and the dependent variable was the Total Immediate Recall score.

Multiple regression revealed that the independent variables significantly predicted the Total Immediate Recall score on the Alternative Scoring System, which accounted for 58% of the variance in performance, $F(9,43) = 7.68$, $p < .0001$. However, only three significant predictors were identified namely the Block Design subtest, Similarities subtest and the Rey Complex Figure Test Copy score. The Block Design subtest, a measure of visuo-construction and problem solving ability was the best predictor in the Total Immediate Recall score of the Alternative Scoring System, which significantly accounted for 38% of the variance, with the Rey Complex Figure Test Copy score and the Similarities subtests each accounted for an additional 10% of the variance in the performance.

3.6.2. Predictors of the Total Delayed Recall score on the Alternative Scoring System

In the regression analyses of the Alternative Form, the independent variables were the Block Design subtest, Object Assembly subtest, Full Scale IQ score, Performance IQ score, and the Rey Complex Figure Test Delay score, and the dependent variable was the Total Delayed Recall score on the Alternative Scoring System.

Multiple linear regression revealed that the independent variables significantly predicted the Total Delayed Recall score of the Alternative Scoring System, which accounted for 26% of the variance in performance, $F(5,43) = 4.06$, $p < .005$. However, only the Block Design subtest significantly predicted the Total Delayed Recall score that accounted for 25% of the variance in performance, $F(1,43) = 15.2$, $p < .0001$.

3.6.3. Predicting performance on the Original Scoring System using the Alternative Scoring System.

A scatter-plot of the performances on immediate recall on the Alternative Scoring System and Original Scoring System was developed in order to generate an equation to determine the accuracy of the Alternative Form to predict performance on the Original Scoring System.

Figure 3.18 illustrates the Total Immediate Recall performances on the Alternative Scoring System and the Original Scoring System

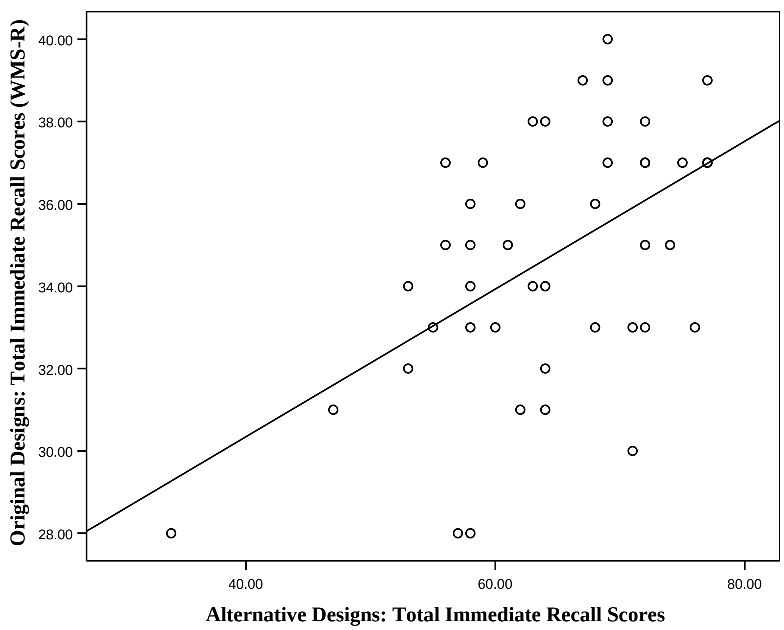


Figure 3.18: The relationship between the Total Immediate Recall scores on the Alternative Scoring System and Immediate Recall performances on the Original Scoring System

Figure 3.18 shows a moderate relationship between the Alternative Scoring System and Original Scoring System is evident. The results of the analysis revealed the following regression equation: Y = 28 + (X x 0.26). However, given that the Alternative Scoring System predicted 25% of the variance in performance on the Original Scoring System, the use of the regression equation would not be an accurate predictive measure for scores between the two scoring systems.

3.7. Additional Memory Measures: Cued and Recognition

In the development of additional measures for the Alternative Form, the hypothesis that the cueing and recognition procedures would provide useful information regarding retrieval processes was investigated. The utility of the additional measures were evaluated using descriptive analyses and paired sample *t*-tests.

3.7.1. Cued Procedure

A cue was provided in this study if a participant could not recall any part of a particular design following delayed recall. Over half of the participants were provided with a cue. A total of 30 participants (68%) were provided with a cue for at least one design and 14 participants (32%) required no cue to facilitate recall of any of the designs as shown in Table 3.35.

Table 3.35: The provision cues to facilitate recall of the Alternative Designs (n = 30).

Number of Cues	Number of Participants	Percentage
Did not require a cue	14	32%
Required at least one cue	30	68%
1 cue	24	55%
2 cues	5	11%
3 cues	1	2%
4 cues	0	0%

Of the 30 participants (68%) that were given a cue, 24 participants (55%) required a cue only for one design, five participants (11%) required a cue for two designs, and one participant (2%) required a cue to assist in the recall of three designs. No participant required a cue to recall all four designs.

Table 3.36: The number of cues given for the Alternative Designs and those who benefited from a cue (n = 30).

Design	Number of participants given cues	Number of participants who benefited from cue	Percentage who benefited
A	13	8	62%
B	19	15	79%
C	–	–	–
D	5	4	80%

Table 3.36 demonstrates that thirteen participants required a cue for Design A and eight benefited from the provision of cue (62%); 19 required a cue for Design B and 15 participants benefited (79%); no participant required a cue for Design C and five participants required a cue for Design D with four (80%) who benefited from being given the cue. In total, 74% of the participants who were provided with a cue benefited from it.

Table 3.37: The mean gain scores on the Alternative Designs of the Alternative Form (n = 30).

Design	n	*M*	*SD*	Median	Range
A	13	3.69	5.42	2	0 – 18
B	19	12.47	7.16	15	0 – 20
C	–	–	–	–	–
D	5	3.20	2.39	3	0 – 6

The gain in scores on each of the four designs is displayed in Table 3.37 indicates that most cues were required for Design B with the greatest gains in scores 12.47, *SD* = 7.16, and the range of scores 0 – 20 points, were obtained on this design. Similarly, Design A obtained a wide range 0 – 18 of performances, but the mean gain 3.69, *SD* = 5.42, and median score of 2 was much smaller than Design B (median = 15). No cues were required for Design C with all participants recalling at least some aspect of the design. Design D had a similar mean of 3.20, *SD* = 2.39, and median score of 3 to Design A, but the range of scores was more restricted from zero to six points.

The results from Table 3.36 and 3.37 indicate that for the designs where no free recall was obtained, the provision of a specific cue facilitated partial or full recall of the particular design for the majority of participants.

Table 3.38: The mean gain in scores on Total Delayed Recall on Alternative Form (n = 30).

Alternative Scoring System	*M*	*SD*	Range
Total Delayed Recall Score	38.3	9.60	16 – 52
Total Delayed Cued Score	48.3	13.62	23 – 72
Total Delayed Score Gain	10.0	8.91	0 – 33

The mean Total Delayed Recall score of the 30 participants who required a cue was 38.3, *SD* = 9.60, with a range of scores from 16 – 52 points as illustrated in Table 3.44. The provision of a cue resulted in a mean Total Delayed Recall to increase to 48.3, *SD* = 13.6, with a range of scores 23 – 72. The overall mean gain in Total Delayed Recall was 10 points, *SD* = 8.91, with a range of score gains from 0 – 33, which was statistically significant, $t(29) = 6.17$, $p > .0005$.

The results indicate that the provision of cues did not only increase the mean Total Delayed Recall score after a delay, but provided a higher bottom score and greater range in the total performances of approximately 13 points (i.e., from 16–52 to 23–72 following a cue). However, the range of total gain scores indicated that some participants did not benefit from a cue. In order to determine the number of participants that did not benefit from at least one cue Figure 3.19 illustrates the distribution of the range of total score gains obtained on the three designs.

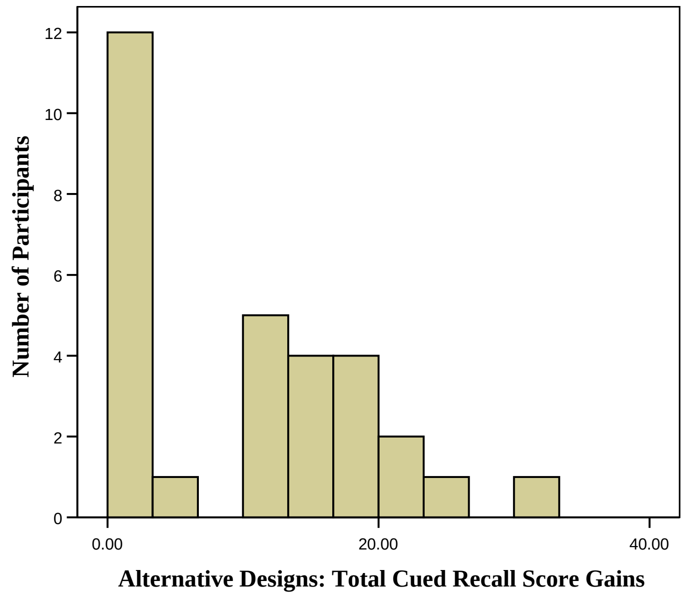


Figure 3.19: The distribution of score gains on the cued recall measure

Figure 3.19 highlights that nine out of 30 participants did not benefit from a cue and three only marginally. Of the 21 participants where partial or full recall of the designs was elicited, the majority obtained a gain of 10 – 20 points. This gain in many cases was approximately 50% of the points potentially allocated to the design.

In order to determine whether an association between benefiting from a cue and general intellectual ability was prevalent, Figure 3.20 illustrates a scatter-plot demonstrating the relationship between these two variables.

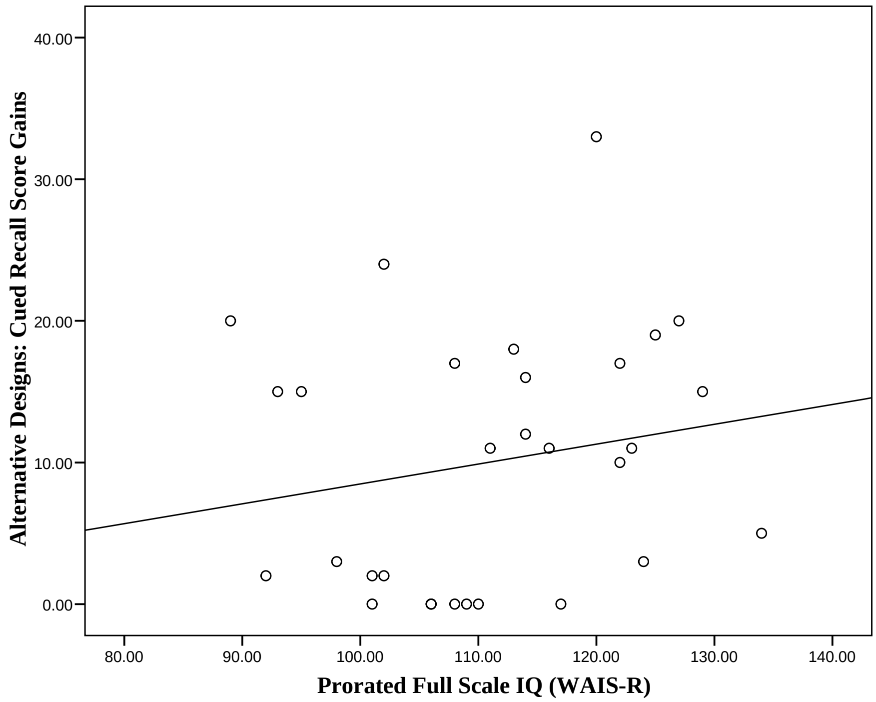


Figure 3.20: The distribution of cued recall score gains and general level of intellectual ability

Figure 3.20 illustrates a positive weak association between score gains and general ability such that participants who made greater gains with the use of a cue did not necessarily demonstrate higher levels of intelligence.

3.7.2. Recognition Procedure

Following delayed free recall each participant was given a multiple choice recognition task for each of the four designs. Participants were administered the recognition procedure regardless of whether the designs were correctly reproduced on delayed recall or whether a benefit from a cue was demonstrated. Table 3.39 illustrates that in the cases where participants did not benefit from a cue, all participants obtained perfect recognition of each of the four designs.

Table 3.39: Performance on the recognition procedure for participants who benefited and did not benefit from a cue (n = 30).

	Percentage who Benefited from a cue	Recognition performance
Design A	62%	100%
Design B	79%	100%
Design C	–	–
Design D	80%	100%
Total Mean	74%	100%

From the 38% of individuals who were provided a cue for Design A, 21% who required a cue for Design B and 20% who required a cue for Design D and did not benefit from it, all were able to correctly identify the target design.

Performance on the recognition measure resulted in a 100% identification rate of each of the four designs for all the participants in this study that were of a non-clinical sample. The results from the recognition procedure indicate that in a sample of individuals with an IQ of 89 to 141 no participant had difficulty in recognising the four designs. From a clinical perspective, failure of a non-clinical sample to recognise the designs would have clinical significance. Therefore, the development of a multiple choice recognition task can potentially provide clinically useful information about memory function beyond which can be obtained from standard administration of only immediate and delayed procedures.

DISCUSSION

The wealth of research and clinical data available on the Visual Reproduction subtest of the Wechsler Memory Scale – Revised makes it a frequently used test in the neuropsychological assessment of memory (Lezak et al, 2004; Moye, 1997).

This study aimed to develop an Alternative Form of the Visual Reproduction subtest of the Wechsler Memory Scale – Revised and additional cued and recognition procedures that have been found helpful for diagnostic purposes in clinical practice. Secondly, this study aimed to develop a scoring system for the Alternative Form that was an improvement on the scoring system of the Visual Reproduction subtest of the Wechsler Memory Scale – Revised. The Alternative Scoring System developed in this study was modelled on a Revised Scoring System developed subtest by Clark (2000) for the Visual Reproduction subtest of Wechsler Memory Scale – Revised. Thirdly, the study aimed to explore the psychometric properties of the Alternative Form and scoring system. This included the reliability and validity of the form by comparing performances of a non-clinical adult population on the Visual Reproduction subtest of the Wechsler Memory Scale – Revised and the Alternative Form.

4.1. Evaluating the Equivalence of the Alternative Form and the Visual Reproduction subtest of the Wechsler Memory Scale – Revised

4.1.1. Design Aspects

In order to maximise the chance of the new version being a good Alternative Form to the Visual Reproduction subtest of the Wechsler Memory Scale – Revised, the development of the designs had to be comparable. As such, four designs were developed in this study to match the four design structure of the Visual Reproduction subtest of the Wechsler Memory Scale – Revised.

Careful consideration was given to the structure and the features of the Alternative Designs. The Alternative Designs of the Alternative Form were carefully developed not to constitute a part of or be closely related to a primary element from another design, so to limit the potential for overlap or confusion between designs.

Each alternative design was individually matched to correspond as closely as possible to the features, size and level of difficulty to the designs of the Visual Reproduction subtest of the Wechsler Memory Scale – Revised. Design A of the Alternative Form consisted of lines and squares as its main features as did Card A of the Visual Reproduction subtest. Design A was developed to be more spatially demanding than Design B. Design B of the Alternative Form used a single regular shape to represent the design as did Card B of the Visual Reproduction subtest of the Wechsler Memory Scale – Revised. Design A and Design B of the Alternative Form were developed to be simpler in structure than Designs C and D. Design C was developed to be easier than Design D with only one large key figure, like that of Card C of the Visual Reproduction subtest. Design C had greater availability to verbal encoding than Design A, Design B and Design D, such that it resembled a pizza. Design D had the most number of spatial elements than Design A, Design B and Design C. Design D represented the same component structure as Card D of the Visual Reproduction subtest of the Wechsler Memory Scale – Revised, with the use of three primary figures.

4.1.2. Administration Procedures

The administration procedure of the Alternative Form was exactly the same to the administration protocol of the Visual Reproduction subtest of the Wechsler Memory Scale – Revised described in the manual (Wechsler, 1987). Specifically, each design was presented for 10 seconds followed by immediate recall of each design. A delayed free recall procedure was administered 25 – 35 minutes after the initial presentation.

4.1.3. Design Comparison of the Alternative Form and the Visual Reproduction subtest

Due to variability of other factors that could potentially influence test performance, the Visual Reproduction subtest of the Wechsler Memory Scale – Revised and the Alternative Form were administered on two separate sessions. This was done to reduce interference and confusion of the two forms and not overload memory. The administration of the Visual Reproduction subtest of the Wechsler Memory Scale – Revised and the Alternative Form occurred over two sessions scheduled 7-10 days apart. The timeline of approximately one week between the administration of the Visual Reproduction subtest and the Alternative Form was considered sufficient not to cause interference effects, with no participants substituting the designs between sessions.

The Visual Reproduction subtest and the Alternative Form were administered to participants in a serial counterbalanced order. This was done so the potential gain in scores from familiarity with the assessment process would be counteracted and test order effects minimised.

Furthermore, the Visual Reproduction subtest and the Alternative Form were administered amongst other cognitive tests commonly used in clinical neuropsychological assessments. That is, the two versions were alternated with other verbal and non-verbal tasks (Logical Memory, Rey Auditory Verbal Learning Test, Rey Complex Figure Test) and a measure of intelligence (Wechsler Adult Intelligence Scale – Revised). This was done to not overload one cognitive domain and to maintain participant interest and motivation.

4.1.4. Controlling for Unknown Factors

To limit the influence of unknown factors such as stress and anxiety on test performance, the stability of performances between Session One and Session Two was evaluated using the Symbol Search and the Speed and Comprehension of Language Processing test. The speed of processing score in Session One set the benchmark in order for second session to proceed, which was met by all participants.

Due to the well documented effects of depression on cognitive functioning, all participants were screened for depression using the Beck Depression Inventory – Second Edition. The Beck Depression Inventory – Second Edition was administered at the end of the second session so questions that could potentially trigger an emotional response did not contaminate test performance. Only one participant was excluded from this study on the basis of a mood disorder.

4.1.5. The Relationship between the Alternative Form and the Visual Reproduction subtest of the Wechsler Memory Scale – Revised

In order for the Alternative Form and the Visual Reproduction subtest of the Wechsler Memory Scale – Revised to be equivalent, a moderate to high correlation between the two versions would be expected. The hypothesis that the Alternative Form would have a moderate-high positive correlation with the Visual Reproduction subtest of the Wechsler Memory Scale – Revised was supported.

The relationship between the Alternative Form and the Visual Reproduction subtest revealed a moderate association on immediate recall (r = .51; r_s = .45) and delayed recall (r = .43; r_s = .37).

To evaluate the robustness of the correlation between the two forms, the stability coefficient for immediate and delayed recall on the Visual Reproduction subtest would set the benchmark to make comparisons. On the Visual Reproduction subtest of the

Wechsler Memory Scale – Revised for participants aged 20 – 64 the stability coefficients ranged from .56 to .80, with an average of .68 on the immediate recall, and a range of .58 to .68 with an average of .63 on delayed recall for a 4 – 6 week test-retest period (Wechsler, 1987). Taking into consideration the stability coefficient of the Visual Reproduction subtest of the Wechsler Memory Scale – Revised, the correlation between the Alternative Scoring System and the Original Scoring System suggests a high association between the Alternative Form and the Visual Reproduction subtest was demonstrated. Therefore, the results of this study indicated that the Alternative Form and the Visual Reproduction subtest of the Wechsler Memory Scale – Revised were highly comparable.

The common variance shared between the Alternative Form and the Visual Reproduction subtest of the Wechsler Memory Scale – Revised for Total Immediate Recall and Total Delayed Recall was 26% and 18% respectively. Furthermore, the Alternative Form predicted 20% and 17% in the variance of scores on the Visual Reproduction subtest of the Wechsler Memory Scale – Revised. However, given the small sample size (n = 44) and the truncated IQ range (89 – 141) in this study, the true correlation between the Alternative Form and the Visual Reproduction subtest of the Wechsler Memory Scale – Revised may be higher if a wider IQ range and larger sample size was obtained, potentially making the two versions even more equivalent.

4.2. The Advantages of the Alternative Scoring System

4.2.1. Design Aspects

The Alternative Scoring System was developed as an improvement to the Original Scoring System of the Visual Reproduction subtest of Wechsler Memory Scale – Revised. The Alternative Scoring System was modelled on the research by Clark (2000) who revised the Original Scoring System of the Visual Reproduction subtest to address the identified problems and shortcomings of the subtest.

The Original Scoring System of the Visual Reproduction subtest was criticised for having differential weighting of each design. The first three designs contributed to 54% to the total score (Clark, 2000). Consequently, undue importance was placed on the fourth design. In fact, no recall on Card D could potential result in a performance below the 25^{th} percentile even if perfect scores were obtained on the other three designs. In contrast, the Alternative Scoring System was developed on an equal number of 20 points allocated for each of the four designs. This avoided making arbitrary decisions about the performance, with each design potentially contributing to 25% of total score providing a consistent metric to each design.

The range of scores for the Alternative Scoring System was extended to a maximum of 80 points which was effectively double the range of 0 – 41 scores of the Original Scoring System. The expansion of the test score range has the potential to increase the discrimination between individual performances, reduce the error score, minimise floor and ceiling effects and increase the discriminatory power of the Alternative Form.

The Alternative Scoring System developed criteria that were scored independently so the score on one item was not contingent on the score on a previous item. Unlike in scoring the designs of the Visual Reproduction subtest of the Wechsler Memory Scale – Revised, a failure on one criterion did not preclude scoring on later criteria on the Alternative Scoring System. Therefore, no assumptions were made about what aspects of the each design were important to be remembered.

The Alternative Scoring System awarded credit for partial recall of designs. This had the potential to better discriminate between poorer performances and reduce floor effects, particularly in the elderly and clinical populations. Additional criteria for perfect reproductions were also included to potentially better discriminate between good performances and reduce ceiling effects, particularly in mild brain injured populations and higher functioning individuals.

The Original Scoring System was criticised for placing undue emphasis on drawing precision and attention to detail, rather than memory ability (Gfeller et al, 1995; Haut et al, 1996; Haut et al, 1994; Heilbronner, 1992). As such, tolerances for carelessness, impulsivity, and poor motor control were included in the Alternative Scoring System. Moreover variations in the tolerances were included to potentially enhance the discrimination between poor and good reproductions. That is, generous tolerances were included for earlier criteria by not allowing carelessness and impulsivity to discount memory performance. More stringent tolerances were included for later criteria on each design to account for the aspect of precision on memory function. The inclusion of tolerances potentially eliminated subjective judgement by having operational criteria.

Another criticism of the Original Scoring System was the written criteria provided greater availability of verbal mediation to confound performance (Loring, 1989). In the development of the Alternative Scoring System the verbal nature of the criteria was acknowledged. However, the scoring criteria were devised to focus more on the spatial aspects and relationships between elements of each design.

4.2.2. Increased Score Range and Score Distributions

A major feature and advantage in the development of the Alternative Scoring System was the increased number of criteria from 41 on the Original Scoring System to 80 on the Alternative Form.

The results of this study indicated that the Alternative Form had a larger distribution of scores than on the Visual Reproduction subtest of the Wechsler Memory Scale – Revised. The score distributions were towards the upper end of the potential range of scores on immediate (M = 64.2, SD = 8.84) and delayed recall (M = 44.8, SD = 13.9) of the Alternative Form. Similarly the participants in this study performed well on immediate recall (M = 34.7, SD = 3.09) and delayed recall (M = 31.8, SD = 4.79) on the Original Scoring System. However, a greater distribution of scores was evident on the Alternative Scoring System for immediate recall (34 – 77) and delayed recall (16 – 80), than on immediate recall (28 – 40) and delayed recall (16 – 39) on the Visual Reproduction subtest. Indeed, not only was there a larger distribution of scores on the Alternative Form than the Visual Reproduction subtest of the Wechsler Memory Scale – Revised, the scores were more normally distributed particularly on delayed recall in this study.

The larger distribution of scores on the Alternative Form has the potential to better discriminate individual performances at higher levels of ability without performances being complicated by ceiling effects. Although ceiling levels were approached on immediate and delayed recall, only one participant obtained a perfect score on the Alternative Form after a time delay. As the participants in this sample were 44 high functioning individuals with an IQ range of 89 – 141 and only one obtaining a perfect score ceiling effects were not evident on the Alternative Scoring System.

The increased number of criteria in the Alternative Scoring System did not indicate a larger range of scores would be obtained. However, in this study the range of scores was substantially greater on the Alternative Form than on the Visual Reproduction subtest. Indeed, the immediate recall scores ranged from 43 points to 64 points on delayed recall on the Alternative Form. In comparison, the score range in this study was 12 points on immediate recall and 23 points on delayed recall on the original Visual Reproduction subtest. Moreover, as delayed recall is considered as the most useful indicator of memory ability, the Alternative Form potentially provides more scope for detection of changes in memory function over time.

The wider range of scores obtained on the Alternative Form in this study also has the potential to minimise floor effects. The results of this study found no floor effects on the immediate recall and delayed recall scores on the Alternative Scoring System. The range of scores below the mean for immediate recall was 64 points and 44 points for delayed recall scores on the Alternative Scoring System. Similarly, no floor effects were obtained on the Original Scoring System in this study, although a smaller range of scores below the mean for immediate recall of 34 points and 31 points for delayed recall were obtained. This suggested that the Alternative Scoring System provided over half the score range for performances at lower levels, potentially increasing the discriminatory power of the Alternative Form. It is particularly important for a test to be able to discriminate between poorer performances as memory difficulties are prominent in many neurological and neuropsychological disorders and are typically a diagnostic feature of dementias and amnesic syndromes (Kopelman, 2002; Papanicolaou, 2006).

Having no floor and ceiling effects evident in this study reflected positively on the Alternative Scoring System as a potentially useful diagnostic instrument. Floor and ceiling effects are particularly problematic with elderly and similarly with mild-brain injured or high functioning individuals respectively (Lezak et al, 2004). Furthermore, floor and ceiling effects can not only influence diagnostic issues, but impact on interpretation of the effectiveness of treatments and intervention, and the rate or extent of recovery (Tulsky et al, 2003). However, future research would be required to quantify the utility of the Alternative Form at lower levels of general ability and in a variety of clinical populations.

Taken together, the larger range and more normal distribution of scores on the Alternative Scoring System to the Original Scoring System, particularly on delayed recall reflects positively on the development of the scoring criteria for the Alternative Form in this study. The expansion of the test score range of the Alternative Form not only has the potential to increase the discrimination between individual performances, reduce the error score, minimise floor and ceiling effects, but can potentially increase the discriminatory power at higher and lower levels of intellectual ability.

4.2.3. Additional Memory Procedures: Cued and Recognition

Cueing and recognition procedures were developed in this study to supplement the Alternative Form. These procedures were developed to provide further clinical and diagnostic information regarding the nature and extent of an underlying memory impairment not provided by the use of free recall alone. Specifically, the value of providing cueing and recognition procedures in determining whether memory difficulties were primarily an encoding, storage or retrieval deficit has been demonstrated (Gass, 1995; Fastenau, 1996; Lezak et al, 2004).

A cue was used to determine whether minimal information was sufficient to trigger memory recall of a particular design. Following the cued procedure if failure to recall the designs occurred, the recognition procedure would be administered to determine if a much richer presentation of the design in the context of other distracters could facilitate recall. The hypothesis that the development of cuing and recognition procedures would provide additional information regarding memory performance was supported.

4.2.3.1. Cueing Procedure

A cued recall procedure was used in this study due to the clinical observation that information previously difficult to recall was facilitated when a cue was provided. The provision of a cue was considered an intermediate step between free recall and recognition. A cue was administered to determine whether the provision of a partial aspect of the design was sufficient to trigger recall without the need for presenting the entire design.

The cues developed in this study were chosen based on a partial feature of each of the designs that was not an obscure or trivial aspect of the design. The cues were carefully selected so that each cue could not be representative of any other design. For example, a square was the best available cue for Design A as a straight line could account for any of the other three designs. This made it easier to identify which cue was administered for a particular design and to determine how much recall was facilitated by the provision of a particular cue. Moreover, as only a partial aspect of a design was provided the cue procedure would only provide a minimal amount of information. The use of a partial cue also would not confound performance on the recognition task, which is important for determining the nature of a memory problem. The cues were administered in a standardised manner for all participants but were only provided if no part of a particular design was generated by free recall, thus making the cued procedure time efficient.

The results in this study indicated that the provision of cues greatly assisted the recall of the Alternative Designs. Sixty-eight percent of participants in this study required at least one cue due to failure in the recall of one or more of the designs. For the participants who failed to recall one or more design, 55% percent were given one cue and 13% were provided with two cues. Of the participants that had apparently forgotten a design, 86% required a cue for the first two designs on the Alternative Form. Of the participants that were provided with a cue, 74% benefited from it. Indeed, a substantial mean gain of 10 points (*SD* = 8.91) in the total score was obtained by the cueing procedure in this sample as a group. The findings suggested that the provision of cues were generally successful in facilitating substantial recall of the designs for participants in this study.

Unfortunately, there was no information in the Wechsler Memory Scale – Revised manual about the number of participants who had entirely forgotten the original Visual Reproduction designs. However, some normative data on this issue was available from the local and stratified norms for the Wechsler Memory Scale – Revised gathered by Shores and Carstairs (2000). Detailed examination of the previous author’s raw data indicated that approximately 22% of the participants forgot one or more design on the Visual Reproduction subtest of the Wechsler Memory Scale – Revised. Of the designs that were forgotten, 95% of the sample that failed to recall the designs was confined to the first two designs (A. E. Shores, R. J. Carstairs., & H. Madill, personal communications, August 12, 2007).

Furthermore, in the study by Clark (2000), 47% of the non-clinical participants required at least one cue on the Visual Reproduction subtest of the Wechsler Memory Scale – Revised. The first two designs were generally “forgotten” more frequently than the other two designs. The provision of a cue was reported to result in gains of more than 50% in the total score on delayed recall (Clark, 2000). Similarly, Gass (1995) found 55% of participants required a cue for at least one design on the Visual Reproduction subtest of the Wechsler Memory Scale – Revised. The previous author reported an average gain of five points with up to an increase of 50% in the total score on delayed recall.

The findings from both this study and the previous research on the Visual Reproduction subtest of the Wechsler Memory Scale – Revised indicated the first two designs were the most difficult to remember on delayed free recall. This reflects favourably for the equivalence of the first two designs on the Alternative Form and the Visual Reproduction subtest of the Wechsler Memory Scale – Revised.

For the majority of participants who required a cue on the Alternative Form benefited from it. However, some participants made substantial gains whereas others did not (range 0 – 20 points). The individuals at higher levels of general ability demonstrated slightly greater gains when a cue was provided to participants at lower IQ range. This may suggest that those with higher intelligence potentially demonstrate a greater ability to use semantic cues to facilitate recall of information shown difficult to retrieve using free recall.

Taken together, information that was previously difficult to recall was facilitated when a cue was provided. The findings in this study provide support for the utility of cues as part of standard administration procedures. Cues should be considered in the standardisation of the Alternative Form in future normative studies. The cueing procedure developed in this study potentially has important implications for the clinical assessment of memory disorders and rehabilitation.

4.2.3.2. Recognition Procedure

A recognition procedure enriches the evaluation of memory ability particularly when poor performance on delayed free recall and when no benefit was achieved from a cue. Recognition measures provide an opportunity to evaluate storage of information without placing additional demands on memory, generating further information of potential clinical significance.

The recognition format in this study was developed to minimise the demands placed on memory by having all distracters and the target design presented on one sheet at one time, but separately for each design. Moreover, the recognition procedure was time efficient with the presentation of only four recognition cards. Recognition memory was assessed with a six alternative forced-choice measure with the target stimulus interspersed amongst five distracters. The choice from six designs allowed for a greater level of confidence in knowing a decision was made based on the recognition of the design rather than a chance performance. The target designs were not placed in the first and last positions on the recognition cards in order to reduce the contribution of chance on performance because the natural bias in guessing tends to favour these positions (Clark, 2000).

In addition, the recognition cards were administered to all participants after delayed recall and following the cued procedure if a cue was given. Therefore, the recognition procedure did not confound immediate and delayed recall performance on the Alternative Form. Furthermore, whether participants that could generate a design on delayed free recall could also correctly recognise it from amongst the distracters was investigated to determine the quality of the procedure.

The development of the distracter designs was devised to retain a similar level of complexity as the target design with a similar number and type of elements. That is, the angles, rotations, order and position of the elements in relation to one another were altered and re-arranged in an attempt to make them distinctly different so not to result in confusion. Conversely, the distracter designs were developed to be similar in shape and size to require accurate recognition of the individual designs and not due to chance.

The results in this study revealed that the presentation of the entire design was correctly identified from a number of distracters by all participants, particularly the participants that did not benefit from a cue. Participants demonstrated traditional patterns of performance expected in a non-clinical sample with Average and above general ability as having no difficulty in recognising the designs. A poor performance on the recognition measure in a non-clinical sample would have clinical significance. The perfect recognition performance validated the quality of the recognition procedure for the Alternative Form and supported the notion that a recognition procedure is a useful measure.

Taken together, the results of the cueing and recognition procedures in this study highlight the importance not to make inferences about memory function solely based on free recall. The routine administration of additional memory procedures following immediate and delayed recall provides an opportunity to potentially obtain clinically important information about the nature and extent of memory problems efficiently. However, the diagnostic application of the cued and recognition procedures developed for the Alternative Form of the Visual Reproduction subtest of the Wechsler Memory Scale – Revised in this study requires further investigation.

4.3. The Psychometric properties of the Alternative Form and Scoring System

4.3.1. Reliability of the Alternative Scoring System

In order for the Alternative Form to serve the purpose in clinical assessment of memory it must have good reliability. The reliability of the Alternative Form would need to be at least equal to the reliability of the Visual Reproduction subtest of the Wechsler Memory Scale – Revised given the smaller sample in this study. The hypothesis that the Alternative Form would demonstrate moderate-high reliability estimates was supported.

The internal consistency of the Alternative Form was estimated using the average of all split half reliability coefficients of each design determined by Cronbach's alpha. The reliability results in this study revealed good internal consistency estimate of .90 for the Total Immediate Recall and .94 for Total Delayed Recall. The reported internal consistency estimates of the Visual Reproduction subtest of the Wechsler Memory Scale – Revised ranged from .46 to .71 on immediate recall and .38 to .59 on delayed recall in the manual (Wechsler, 1987; Williams et al, 1998). The greater number of items for each of these designs likely increased the reliability of the Alternative Form.

The internal consistency estimates for the individual design revealed adequate reliability ranging from .80 to .87 on immediate recall for three of the designs, with estimates even higher on delayed recall, ranging from .91 to .96. Even though the third design demonstrated the lowest internal consistency for both immediate recall (r = .61) and delayed recall (r = .66), the Alternative Scoring System overall indicated better reliability than the Original Scoring System of the Visual Reproduction subtest of the Wechsler Memory Scale – Revised.

However, no information regarding the reliability of the designs of the Visual Reproduction subtest was provided in the Wechsler Memory Scale – Revised manual (Thiesen et al, 1997). Therefore, individual reliability comparisons can not be made between the designs of the Alternative Form and the designs of the Visual Reproduction subtest. This would have been beneficial in order for comparisons to be made between the designs.

This study was conducted by a single researcher that was involved in the development of the scoring system and application of criteria to each of the designs. Due to the researcher being involved in the development of the scoring system for each of the designs and the application of criteria no formal measure of intra-rater reliability was conducted. It would be expected that high intra-rater reliability would have been determined given the familiarity of the researcher with the scoring criteria would not have been useful or representative of the Alternative Scoring System. However, a checking procedure was incorporated in the scoring of the Alternative Designs as a great deal of time and attention was spent on the development of the scoring criteria for each design.

4.3.2. Validity of the Alternative Scoring System

In order for the Alternative Form to be established as an alternative to the Visual Reproduction subtest of the Wechsler Memory Scale - Revised, it needs to not only be reliable but demonstrate that it measures the construct it purports to measure (Cohen, 1988; Tabachnick & Fidell, 1996). The convergent validity of the Alternative Form was investigated using the Rey Complex Figure Test due to its wide clinical use as a measure of non-verbal memory and its sensitivity to brain impairment (Meyers & Meyers, 1995; Lezak et al, 2004; Moye 1997). The hypothesis that the scoring system of the Alternative Form would demonstrate convergent validity with moderate-high positive associations with a non-verbal memory measure, namely the Rey Complex Figure Test was supported.

The Rey Complex Figure Test has reported factor loadings on visuo-constructional and visuo-spatial ability, planning and organisational skills on the copy trial and visuo-spatial and memory factors on delayed recall in the manual (Meyers & Meyers, 1995). Similarly, factor analytic studies have reported immediate recall on the Wechsler Memory Scale – Revised to load on visuo-constructional and perceptual factors and delayed recall to better reflect memory functioning (Bowden et al, 2001; Heilbronner, 1992).

The Alternative Form demonstrated a moderate relationship with the Rey Complex Figure Test (r = .40) on delayed recall. Similarly performance on immediate recall demonstrated a strong relationship with the Copy trial of the Rey Complex Figure Test (r = .61). The meaningful relationships with the Rey Complex Figure Test in this sample supported the convergent validity of the Alternative Form as a measure of non-verbal memory. However, future research on establishing convergent validity of the Alternative Form with other measures of non-verbal memory will need to consider the reliability of the measures compared. Such that, a failure to recognise lower than optimal reliability may attenuate validity correlations and encourage unnecessary proliferation of the underlying construct (Schmidt & Hunter, 1996).

4.3.3. Discriminant Validity

The discriminant validity of the Alternative Form was evaluated using the Logical Memory subtest of the Wechsler Memory Scale – Revised and the Rey Auditory Verbal Learning Test given their wide clinical application and utility in the assessment of verbal memory function (Lezak et al, 2004).

The hypothesis that the scoring system of the Alternative Form would demonstrate discriminant validity with weak associations with other measures of verbal memory and learning was supported. A weak relationship between the Alternative Form and the Logical Memory subtest on immediate recall (r = .19) and delayed recall (r = .27) with no meaningful relationship with the Rey Auditory Verbal Learning Test Learning score (r = .15) and Delayed Recall (r = .07). The weak correlations on the Alternative Scoring System and other verbal measures support for the discriminant validity of the Alternative Form of the Visual Reproduction subtest of the Wechsler Memory Scale – Revised and further support as a measure of <u>non-verbal</u> memory.

Concerns regarding the Visual Reproduction subtest of the Wechsler Memory Scale – Revised as a measure of non-verbal memory have stemmed from inconsistent findings in its ability to distinguish left from right lateralised hemispheric damage (Chelune & Bornstein, 1988; Squire & Butters, 1992; Baxendale, 1997; Squire, 1986; Jones-Gotman, 1986; Chelune & Bornstein, 1988). Indeed, Naugle, Chelune, Cheek, Luders and Awad (1993) reported patients with left lobotomies showed a significant drop in measures of verbal memory, but no decrements in nonverbal memory were found in those with right temporal lobotomy. In fact, there was no indication right temporal lobotomy resulted in reduced performance on immediate or delayed procedures relative to preoperative baseline or to non-surgical intractable patients. However, patterns of non-verbal memory impairments were reported to only emerge as a within subject effect in a sample of a variety of unilateral brain lesions (Chelune & Bornstein, 1988). Furthermore, the Visual Reproduction has also demonstrated utility in distinguishing between amnesic disorders, cortical and subcortical dementias (Butters et al, 1988).

The inconclusive findings in the literature on the Visual Reproduction subtest of the Wechsler Memory Scale – Revised as a measure of non-verbal memory may well reflect methodological issues. Due to the rarity in circumscribed right temporal lobe lesions, the inclusion of participants with a variety of brain-damaged regions may potentially mask any mild deficits in non-verbal memory when evaluating group differences (Larrabee & Curtiss, 1995; Smith et al, 1992). The discriminatory power of the Alternative Form was not investigated in this study. Therefore, the sensitivity and specificity of the Alternative Form in detecting non-verbal memory impairment in a variety of clinical populations is important to be quantified by future research. Furthermore, the application of conservative methodological procedures in future studies is warranted in order for the validity of the Alternative Form to be accurately established.

4.3.3. Construct Validity

Given the well known relationship between memory and intelligence, it was important to establish the construct validity of the Alternative Form as a measure of memory and not general ability with only a weak – moderate relationship with IQ scores.

The hypothesis that the Alternative Form would demonstrate construct validity with weak – moderate positive correlation with the Wechsler Adult Intelligence Scale – Revised, a measure of general ability, was supported. Immediate recall on the Alternative Form demonstrated a moderate relationship with measures of intelligence, namely the Full Scale IQ (r = .40), Performance IQ (r = .39) and Verbal IQ (r = .33) scores of the Wechsler Adult Intelligence Scale – Revised. On delayed recall only a meaningful relationship was evident with the Full Scale IQ (r = .36) and Performance IQ (r = .40) scores. Although some commonality was shared between the Alternative Scoring System and intelligence scores, the relationship was not substantial to suggest it was a measure of general ability rather supported the construct validity of the Alternative Form as a measure of non-verbal memory.

Further analysis of the immediate and delayed performances on the Alternative Form indicated greater associations with non-verbal skills on delayed recall, well known to better represent memory function (Squire, 1986; Wilson, 2004). Specifically, the performances on the Alternative Scoring System shared greatest commonality with measures of whole-part relationships and organisation of visual material (Object Assembly subtest), particularly visuo-construction and non-verbal problem solving ability (Block Design subtest).

4.3.4. Other Findings

It was interesting to note that no meaningful relationship was evident between general intellectual ability and the immediate recall score (*r* = .18) and the delayed recall score (*r* = .22) on the Original Scoring System of the Visual Reproduction subtest of the Wechsler Memory Scale – Revised. The weak insignificant relationship of the Original Scoring System and general ability indicated that on immediate recall the Original Scoring System did not load on general ability. Moreover, the Visual Reproduction subtest may well be evaluating other cognitive skills like memory. Indeed, moderate associations were indicated with other non-verbal skills much like on the Alternative Form.

The moderate association of the Alternative Scoring System and the Visual Reproduction subtest of the Wechsler Memory Scale – Revised with measures of non-verbal ability indicates that the two versions are comparable, particularly on delayed recall in this study. The results in this study provide further support for the equivalence of the Alternative Form and the Visual Reproduction subtest of the Wechsler Memory Scale – Revised.

The reported correlation between immediate and delayed recall of the Visual Reproduction subtest of the Wechsler Memory Scale – Revised ranged from .36 to .67 in the manual, with .67 for the 35 – 44 age group of this study (Wechsler, 1987). The correlation of immediate recall and delayed recall scores on the Visual Reproduction subtest was moderate (*r* = .38), with 14% common variance in scores for the sample of this study. The lower correlation between immediate and delayed recall performances on the Visual Reproduction subtest of the Wechsler Memory Scale – Revised in this study compared to the correlations reported in the manual was potentially an artefact of the small sample size in this study.

In contrast, a strong correlation (r = .53) was found between the immediate recall and delayed recall scores of the Alternative Form in this sample, with 28% of common variance in scores. In making comparisons between the immediate and delayed recall correlations of the two forms, variability in the characteristics of the sample was controlled for as the same sample was used for both analyses. The higher correlation of the immediate recall and delayed recall scores on the Alternative Form than on the Visual Reproduction subtest reflected positively on the Alternative Form as potentially having better psychometric properties than the Visual Reproduction subtest.

The results of this study did not reveal significant age and education effects although there was a general trend for performance on the Alternative Form to decrease as a function of age and increase with higher years of education, particularly for delayed recall performances. The results however revealed differences between males and females on immediate recall, with males performing better overall, but this difference between the sexes was not evident on delayed recall.

The results of this study revealed that despite a maximum two-point gain evident on delayed recall score when the Alternative Form was administered in the second session, this gain was not significant. This indicates that test order did not confound immediate or delayed performance on the Alternative Form. This is important if the normative data for the Alternative Form and the other measures administered are extended as a flexible test battery for use in clinical practice.

4.4. Problems and Limitations of the Study

A problem identified in this study was the restricted score range on immediate recall (12 – 20) and delayed recall (7 – 20) for Design C on the Alternative Form. In addition, the relatively high mean score of 17.9 (SD = 1.74) on immediate recall was comparable to the mean score of 17.4 (SD = 2.46) obtained on delayed recall on Design C, which was much higher than the other three designs on the Alternative Form. Design C was developed to be less spatially demanding than the other designs, thus verbal encoding may have enhanced performance on this design.

Furthermore, the higher performance on Design C may well reflect the truncated IQ range. Given the high functioning of the participants in this study and the greater availability to verbal encoding of the design, this may reflect the ability of the sample as a group to readily make semantic associations between elements of the design. In a sample with a wider IQ range and sample size the performances on the design may be more normally distributed.

No information was available on the third design on the Visual Reproduction subtest in the manual (Wechsler, 1987). Some data regarding performances on the third design was obtained from the local normative study conducted by Shores and Carstairs (2000). Detailed examination of the raw data indicated that 80% of the 399 participants in the study achieved a score of 6 – 9 on the third design. This indicates that the performances on the third designs on both the Alternative Form and the Visual Reproduction subtest of the Wechsler Memory Scale – Revised were performed towards the upper end of the potential range of scores. Although the restricted range of scores potentially reduces the discriminatory power and increases ceiling effects, the results provide support for the equivalence of the Alternative Designs to the designs of the Visual Reproduction subtest of the Wechsler Memory Scale – Revised.

Moreover, the internal reliability coefficient for Design C of the Alternative Form of .61 on immediate recall and .66 on delayed was inadequate. This suggested that as participants generally scored on most items the failure to score on a few items can significantly impact on the internal reliability of the design. Furthermore, the scoring criteria may be too lenient for Design C on the Alternative Form. The high scores on Design C of the Alternative Form may also reflect that the scoring criteria may have been too lenient for this particular design.

In order to remedy the problem of the restricted range of scores, high mean performance and inadequate internal consistency estimate for Design C of the Alternative Form, the scoring criteria may require revision and further refinement. That is, the criteria of the design that have a very high correlation with each other can be eliminated as these items likely measure “the same thing.” Eliminating five items and replacing them with harder criteria may also be effective in providing a greater distribution and range of scores on the design.

The representativeness of the American based norms of the Visual Reproduction subtest Wechsler Memory Scale – Revised in a geographically and culturally unique composition of the Australian population warrants caution in the interpretation of test performance. Furthermore, the application of the Visual Reproduction subtest normative data to the sample in this study needs to be carefully interpreted. That is, the interpolated normative data of the Wechsler Memory Scale – Revised for three of the age bands and specifically the 35 – 44 year olds has been criticised for overestimating performances at upper end of the distribution on the Visual Memory and Delayed Recall Indices that includes the Visual Reproduction subtest (D’Elia et al, 1989; Mittenberg & Burton, 1992).

Moreover, the interpretation of performances on the original Visual Reproduction subtest may also be misleading as the sample sizes of the Wechsler Memory Scale – Revised were small with around 50 to 55 participants in each of the six age bands were used to generate the normative data. The statistical power from a sample of 44 participants was considered sufficient to draw meaningful conclusions from the data, but the sample size in this study was relatively small (Cohen, 1988; Tabachnick & Fidell, 1996). The sample characteristics limit the generalisations of performances to the general Australian population as this was not a stratified and randomised study. Such that, participants in this study were recruited via convenience sampling which can be potentially misleading and biased towards education, ethnicity and socio-economic status (Holdnack et al, 2004).

The psychometric properties of the Alternative Form and scoring system need to be determined using a wider IQ range, particularly in the lower range of intellectual functioning. Furthermore, a larger sample size that is randomly recruited and stratified in an Australian population across a wider age range would also be required. In particular, further studies in the development of a wider normative base with substantial numbers particularly for the older age groups (55 years and over) would be advantageous. It is important to establish representative and extensive data in the older age groups as often in these later years of life memory problems are initially identified and neuro-degenerative diseases are diagnosed.

A disadvantage of this study was that no inter-rater reliability data was available. It would be valuable to obtain an indication of the inter-rater reliability between experienced clinicians and novice scorers to assist with the refinement of the Alternative Scoring System. The scoring system for the Alternative Form was developed and scored by a single researcher with considerable exposure to individual items and their interpretation. As such, the immediate and delayed recall performances on the Alternative Form may have been different if a researcher who was naïve to the development of the design and scoring system process scored the protocols. However, the explicit nature and operational criteria with the addition of example full credit and no credit drawings to provide further clarity in the scoring of the Alternative Designs.

Another problem identified in this study was the large number of participants (68%) who completely forget at least one of the designs on the Alternative Form. This high rate was considered unusual for a non-clinical population and in some clinical populations. This unexpected high rate of “forgetting” may well have reflected the lack of experience of the researcher in the clinical administration of neuropsychological tests and clinical experience at the time of data collection. Indeed, the participants in this study demonstrated that the designs were not necessarily forgotten on the cued procedure.

Similarly, in the research by Clark (2000) a high percentage of non-clinical participants (47%) failed to recall at least one design. Moreover, given a substantial number of participants benefited with the provision of cue that likely reflected the lack of clinical experience of the researcher in test administration rather than it representing a normal phenomenon. Furthermore, the need for prompts, encouragement and reflection time may well be sufficient to facilitate recall of the designs. Standardised guidelines clearly need to be developed to address this issue not only for future studies using the Alternative Form, but for all test manuals.

4.5. Future Directions

Future studies can use contemporary psychometric analyses in determining the equivalence of the Alternate Form. Specifically, item response theory (IRT) and latent variable modelling techniques can be employed (Brown, 2006).

However, the application of item response theory requires large sample sizes of approximately 500 to 1000 participants to adequately obtain an estimate of item difficulty (Osterlind, 2006). In addition, the assumptions that underlie item response theory, such as local independence are not usually met, limiting the application of IRT to the clinical assessment of memory (Osterlind, 2006). Consequently, this technique for item analysis was beyond the scope of the current study.

The empirical strategy of item level latent variable modelling is manifestly the appropriate direction for future studies on the Alternative Form. Item level confirmatory factor analysis will allow delineation of underlying test item constructs and the examination of equivalence, development of scaled scores and subsequent normative research if appropriate (Little, Cummingham, Shahar & Widaman, 2002).

On a theoretical level, the use of latent variable modelling in future empirical studies on the Visual Reproduction subtest will likely add an appropriate element to the perspective of cognitive ability (McGrew, 1997).

A criticism of the non-verbal memory measures such as the Visual Reproduction subtest of the Wechsler Memory Scale – Revised is that a motor response is required. Thus, impairments in constructional or visuo-motor ability can potentially confound performance. Even though the scoring system was explicit and developed to include tolerances in the cases where drawing difficulties are evident, the influences of constructional difficulties on performance are less well controlled. The advantage of establishing the utility of the recognition procedure in a variety of clinical groups can provide information about the contribution of constructional deficits to design recall. Also the development of copy procedure to adjust for motor contributions may be an aim for future research.

In acknowledging that verbal processes are likely employed when processing non-verbal information, the development of a non-verbal index has the potential to identify participants with primarily non-verbal memory difficulties. An index of scoring items that are performed poorly by individuals with non-verbal memory deficits could be developed to potentially discriminate between clinical populations and quantify the severity of memory impairment.

4.6. Conclusion

The aim of this study was to develop an alternative form of the Visual Reproduction subtest of the Wechsler Memory Scale – Revised. A second aim of this study was to develop a scoring system for the Alternative Form. The Alternative Form generated a wide range and distribution of scores in a non-clinical adult population with average and above general intellectual skills, with adequate scope for performances at lower levels of functioning. The use of cue and recognition measures provided useful information that enhanced the diagnostic utility of the Alternative Form.

Preliminary analysis revealed the Alternative Form demonstrated at least as good reliability as the Visual Reproduction subtest of the Wechsler Memory Scale – Revised. The Alternative Form developed in this study had a good correlation with the Visual Reproduction subtest, was correlated with other tests of non-verbal ability and did not have high correlations with other verbal memory measures.

This research constituted a preliminary study of the reliability and validity of an Alternative Form for the Visual Reproduction subtest of the Wechsler Memory Scale – Revised with respect to a modest sample of a non-clinical adult population. The findings of this study suggest that the Alternative Form may well be a valuable clinical tool, particularly for times of serial assessments after further refinement and an appropriate normative study.

REFERENCES

Aggleton, J. P., & Brown, M. W. (1999). Episodic memory, amnesia, and the hippocampal-anterior thalamic axis. *Behavioral and Brain Sciences, 22,* 425 – 489.

Atkinson, R. C., & Shiffrin, R. M. (1968). 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* (Vol. 2, pp. 82 – 90). New York: Academic Press.

Baddeley, A. D. (1986). *Working Memory*. Oxford: Oxford University Press.

Baddeley, A. D. (1994). Working Memory: The interface between memory and cognition. In D. L. Schacter & E. Tulving (Eds.), *Memory Systems* (pp. 352 – 368). Cambridge, MA: MIT Press.

Baddeley, A. D. (1995). The psychology of memory. In A. D. Baddeley, B. A. Wilson, & F. N. Watts (Eds.), *Handbook of Memory Disorders.* England: John Wiley & Sons Ltd.

Baddeley, A. D. (1997). *Human Memory: Theory and practice*. (2^{nd} ed.). Hove: Psychology Press.

Baddeley, A. D. (1999). *Essentials of Human Memory*. London: Psychology Press.

Baddeley, A. D. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Science, 4,* 417 – 423.

Baddeley, A. D. (2004). The psychology of memory. In A. Baddeley, M. D. Kopelman, & B. A Wilson (Eds.), *The Essential Handbook of Memory Disorders for Clinicians* (pp. 1 – 13). England: John Wiley & Sons Ltd.

Baddeley, A. D., Emslie, H., & Nimmo-Smith, I. (1992). *The Speed and Capacity of Language Processing Test Manual.* Bury St Edmunds: Thames Valley Test Co.

Baddeley, A. D., Hitch, G.J. (1974). Working memory. In G.A. Bower (Ed.), *The Psychology of Learning and Motivation* (pp. 47 – 89). New York: Academic Press.

Baxendale, S. A. (1997). The role of the hippocampus in recognition memory. *Neuropsychologia, 35,* 591 – 598.

Baxter, M. G., & Murray, E. A. (2001). Impairments in visual discrimination learning and recognition memory produced by neurotoxic lesions of rhinal cortex in rhesus monkeys. *European Journal of Neuroscience, 13,* 1228 – 1238.

Beck, A. T., Steer, R. A. & Brown, G. K. (1996). *Beck Depression Inventory-II manual.* San Antonio, TX: Psychological Corporation.

Benedict, R. H. B., & Zgaljardic, D. J. (1998). Practice effects during repeat administrations of memory tests with and without alternate forms. *Journal of Clinical and Experimental Neuropsychology, 20,* 339 – 352.

Benton, A. L. (1992). *The Revised Visual Retention Test* (5th ed.). New York: Psychological Corporation.

Bornstein, R. A., & Chelune, G. J. (1988). Factor structure of the Wechsler Memory Scale – Revised. *The Clinical Neuropsychologist, 2,* 107 – 115.

Bornstein, R. A., & Chelune, G. J. (1989). Factor structure of the Wechsler Memory Scale – Revised in relation to age and educational level. *Archives of Clinical Neuropsychology, 4,* 15 – 24.

Bowden, S. C., Carstairs, J. R., & Shores, E. A. (1999). Confirmatory factor analysis of combined Wechsler Adult Intelligence Scale – Revised and Wechsler Memory Scale – Revised scores in a healthy community sample. *Psychological Assessment, 11*, 339-344.

Bowden, S. C., Ritter, A. J., Carstairs, J. R., Shores, A. E., Pead, J., Greeley, J. D., Whelan, G., Long, C. M., & Clifford, C. C. (2001). Factorial invariance for combined Wechsler Adult Intelligence Scale – Revised and Wechsler Memory Scale – Revised scores in a sample of clients with alcohol dependency. *The Clinical Neuropsychologist, 15(1)*, 69 – 80.

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

Broadbent, D. E. (1958). *Perception and Communication*. New York: Pergamon Press.

Broadbent, N. J., Clark, R. E., Zola S., & Squire, L. R. (2002). The medial temporal lobe and memory. In L. R. Squire & D. L. Schacter (Eds.), *Neuropsychology of Memory* (pp. 3 – 23). New York: Guilford Press.

Brown, T. A. (2006). *Confirmatory factor analysis for applied research.* New York, US: Guilford Press.

Buckner, R. L., Kelley, W. M., & Petersen, S. E. (1999). Frontal cortex contributes to human memory formation. *Nature Neuroscience, 2,* 311 – 314.

Burton, B. D., Mittenberg, W., & Burton, C. A. (1993). Confirmatory factor analysis of the Wechsler Memory Scale – Revised standardization sample. *Archives of Clinical Neuropsychology, 8*, 467 – 475.

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

Butters, N., Delis, D. C., & Lucas, J. A. (1995). Clinical assessment of memory disorders in amnesia and dementia. *Annual Review of Psychology, 46,* 493 – 524.

Butters, N., Salmon, D. P., Cullum, M. C., Cairns, P., Troster, A. I., Jacobs, D., Moss, M., & Cermak, L. S. (1988). Differentiation of amnesic and demented patients with the Wechsler Memory Scale – Revised. *The Clinical Neuropsychologist, 2,* 133 – 148.

Carr, A. C. (1982). Memory deficit after fornix section. *Neuropsychologia*, *20,* 95 – 98.

Carstairs, J. R., & Shores, A. E. (2000). The Macquarie University Neuropsychological Normative Study (MUNNS): Rationale and methodology. *Australian Psychologist, 35,* 36 – 40.

Cauthen, N. R. (1977). Extension of the Wechsler Memory Scale norms to older age groups. *Journal of Clinical Psychology, 85,* 181 – 200.

Cermak, L.S. (1984). The episodic-semantic distinction in amnesia. In L. R. Squire & N. Butters (Eds.), *Neuropsychology of Memory*. New York: Guilford Press.

Chelune, G. J., & Bornstein, R. A. (1988). WMS – R patterns among patients with unilateral brain lesions. *The Clinical Neuropsychologist, 2,* 121 – 132.

Cipolotti, L., Shallice, T., Chan, D., Fox, N., Scahill, R., Harrison, G., Stevens, J., & Rudge, P. (2001). Long-term retrograde amnesia….the crucial role of the hippocampus. *Neuropsychologia, 39,* 151 – 172.

Clark, S. E., & Shiffrin, R. M. (1992). Cuing effects and associative information in recognition memory. *Memory & Cognition, 20,* 580 – 598.

Clark, T. (2000). *Development of an Alternative Scoring System for the Visual Reproduction subtest of the Wechsler Memory Scale – Revised.* Unpublished doctoral dissertation, Victoria University, Australia.

Cohen, N. J. (1988). *Statistical power analysis for the behavioural sciences.* Hillsdale, New Jersey: Erlbaum.

Cohen, N. J., & Squire, L. R. (1980). Preserved learning and retention of pattern analyzing skill in amnesia: Dissociation of knowing how and knowing what. *Science, 210,* 207 – 209.

Cornoldi, C., & Vecchi, T. (2003). *Visuo-spatial working memory and individual differences.* New York: Psychology Press.

Corkin, S., Amaral, D. G., Gonzalez, R. G., Johnson, K. A., & Hyman, B. T. (1997). H.M.'s medial temporal lobe lesion: Findings from magnetic resonance imaging. *Journal of Neuroscience, 17,* 3964 – 3979.

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

Delis, D. C. (1989). Neuropsychological assessment of learning and memory. In F. Boller. & J. Grafman (Eds.), *Handbook of Neuropsychology* (pp. 3 – 33). Amsterdam: Elsevier.

Delis, D. C., Kramer, J. H., Kaplan, E., & Ober, B. A. (2000). *California Verbal Learning Test: Second Revision.* San Antonio, TX: The Psychological Corporation.

Delis, D. C., Kramer, J. H., Kaplan, E., & Ober, B. A. (1987). *California Verbal Learning Test: Adult Version.* San Antonio, TX: The Psychological Corporation.

De Lemos, M. (1981). *WAIS-R Australian supplement.* Hawthorn: Australian Council for Educational Research.

Deweer, B, Pillon, B, Pochon, J. B., & Dubois, B. (2001). Is the H. M. story only a 'remote memory'? Some facts about hippocampus and memory in humans. *Behavioural Brain Research*, *127*, 209 – 224.

Dolan, R. J., & Fletcher, P. C. (1997). Dissociating prefrontal and hippocampal function in episodic memory encoding. *Nature, 388*, 582 – 585.

Doss, R. C., Chelune, G. J., & Naugle, R. I. (2000). Comparability of the expanded WMS – III standardization protocol to the published WMS – III among right and left temporal lobectomy patients. *The Clinical Neuropsychologist, 14*, 468 – 473.

Dowling, P. (1998). *The new editions of the Wechsler Adult Intelligence Scale (III) and the Wechsler Memory Scale (III): How will the revisions contained in these tests impact on neuropsychological practice?* Workshop paper presented at the APS College of Clinical Neuropsychologists National Conference, October 8-11, Lorne, Victoria, Australia.

Eadie, K., & Shum, D. (1995). Assessment of visual memory: A comparison of Chinese characters and geometric figures as stimulus materials. *Journal of Clinical and Experimental Neuropsychology, 17,* 731 – 739.

Elwood, R. W. (1991). Factor structure of the Wechsler Memory Scale – Revised (WMS-R) in a clinical sample: A methodological reappraisal. *The Clinical Neuropsychologist, 5*, 329 – 337.

Emilien, G., Durlach, C., Antoniadis, E., Van der Linden, M., & Maloteaux, J-M. (2004). *Memory: Neuropsychological, imaging and psychopharmacological perspectives.* New York: Psychology Press.

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

Eslinger, P. J. (2002). *Neuropsychological interventions: Clinical research and practice.* USA: The Guilford Press.

Fahle, M., & Daum, I. (1997). Visual learning and memory as functions of age. *Neuropsychologia, 35*, 1583 – 1589.

Fastenau, P. S. (1996). An elaborated administration of the Wechsler Memory Scale – Revised. *The Clinical Neuropsychologist, 10,* 425 – 434.

Fletcher, P. C., Shallice, T., & Dolan, R. J. (1998). The functional roles of prefrontal cortex in episodic memory I: Encoding. *Brain, 121*, 1239 – 1248.

Fletcher, P. C., Shallice, T., Frith, C. D., Frackowiak, R. S. J., & Dolan, R. J. (1998). The functional roles of prefrontal cortex in episodic memory II: Retrieval. *Brain, 121*, 1249 – 1256.

Fowler, R. S. (1969). A simple non-language test of new learning. *Perceptual and Motor Skills, 29,* 895 – 901.

Frisk, V., & Milner, B. (1990). The role of the left hippocampal region in the acquisition and retention of story content. *Neuropsychologia, 28,* 349 – 359.

Gaffan, D., Parker, A., & Easton, A. (2001). Dense amnesia in the monkey after transaction of fornix, amygdale and anterior temporal stem. *Neuropsychologia, 39*, 51 – 70.

Gass, C. S. (1995). A procedure for assessing storage and retrieval on the Wechsler Memory Scale – Revised. *Archives of Clinical Neuropsychology, 10,* 475 – 487.

Gfeller, J. D., Meldrum, D. L., & Jacobi, K. A. (1995). The impact of constructional impairments on the WMS – R visual reproduction subtest. *Journal of Clinical Psychology, 51*, 58 – 63.

Gladsjo, J.A., McAdams, L.A., Palmer, B. W., Moore, D.J., Jeste, D.V., & Heaton, R.K. (2004). A six-factor model of cognition in schizophrenia and related psychotic disorders: Relationships with clinical symptoms and functional capacity. *Schizophrenia Bulletin, 30*(4), 739-754.

Gleisnner, U., Helmstaedter, C., Schramm, J., & Elger, C. E. (2002). Memory outcome after selective amygdalohippocampectomy: A study in 140 patients with Temporal Lobe Epilepsy. *Epilepsia, 43*, 87 – 95.

Glosser, G., Goodglass, H., & Biber, C. (1989). Assessing visual memory disorders. *Psychological Assessment, 1,* 82 – 91.

Gooding, P. A., Isaac, C. L., & Mayes, A. R. (2005). Prose recall and amnesia: More implications for the episodic buffer. *Neuropsychologia*, *43*, 583 – 587.

Graf, P., & Schacter, D. L. (1985). Implicit and explicit memory for new associations in normal subjects and amnesic patients. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *11*, 501 – 518.

Graham, K. S., Simons, J. S., Pratt, K. H., Patterson, K., & Hodges, J. R. (2000). Insights from semantic dementia on the relationship between episodic and semantic dementia. *Neuropsychologia, 38*, 313 – 324.

Gravetter, F. J., & Wallnau, L. B. (1996). *Statistics for the Behavioral Sciences* (4th ed.). USA: West Publishing Company.

Hannay, H. J., & Levin, H. S. (1985). Selective Reminding Test: An examination of the equivalence of four forms. *Journal of Clinical and Experimental Neuropsychology, 7*, 251 – 263.

Hannay, H. J., Levin, H. S., & Grossman, R. J. (1979). Impaired recognition memory after head injury. *Cortex, 15*, 269 – 283.

Haut, M. W., Weber, A. M., Demarest, D., Keefover, R. W., & Rankin, E. (1996). Controlling for constructional dysfunction with the visual reproduction subtest of the Wechsler Memory Scale – Revised in Alzheimer's disease. *The Clinical Neuropsychologist, 10*, 309 – 312.

Haut, M. W., Weber, A. M., Wilhelm, K. L., Keefover, R. W., & Rankin, E. D. (1994). The visual reproduction subtest as a measure of visual perceptual/constructional functioning in dementia of the Alzheimer's type. *The Clinical Neuropsychologist*, *8*, 187 – 192.

Hawkins, K. A. (1998). Indicators of brain dysfunction derived from graphic representations of the WAIS-III/WMS-III technical manual clinical samples data: A preliminary approach to clinical utility. *The Clinical Neuropsychologist, 12,* 535 – 551.

Hawkins, K. A., & Tulsky, D. S. (2004). Replacement of the faces subtest by visual reproductions within Wechsler Memory Scale – Third Edition (WMS – III) visual memory indexes: Implications for discrepancy analysis. *Journal of Clinical and Experimental Neuropsychology, 26,* 498 – 510.

Heilbronner, R. L. (1992). The search for a 'pure' visual memory test: Pursuit of perfection? *The Clinical Neuropsychologist, 6,* 105 – 112.

Herman, D. O. (1988). Development of the Wechsler Memory Scale – Revised. *The Clinical Neuropsychologist, 2,* 102 – 106.

Hodges, J. R., Patterson, K., Oxbury, S., & Funnell, E. (1992). Semantic dementia: Progressive fluent aphasia with temporal lobe atrophy. *Brain, 115,* 1783 – 1806.

Hoffman, R. G., Scott, J. G., Tremont, G., Adams, R. L., & Oommen, K. J. (1997). Cross-validation of a method for predicting Wechsler Memory Scale – Revised index scores. *The Clinical Neuropsychologist, 11,* 402 – 406.

Holdnack, J. A., & Delis, D. C. (2004). Parsing the recognition memory components of the WMS – III face memory subtest: Normative data and clinical findings in dementia groups. *Journal of Clinical and Experimental Neuropsychology, 26,* 459 – 483.

Holdnack, J. A., Lissner, D., Bowden., S., & McCarthy, K. A. L. (2004). Utilising the WAIS – III/WMS – III in clinical practice: Update of research and issues relevant to Australian normative research. *Australian Psychologist, 39,* 220 – 227.

Horton, A. M. (1999). Book and test review. *Archives of Clinical Neuropsychology, 14*, 473 – 477.

Howieson, D. B., & Lezak, M. D. (2004). Separating memory from other cognitive disorders. In A. D. Baddeley, M. D. Kopelman & B. A. Wilson (Eds.), *The Essential Handbook of Memory Disorders* (pp. 179 – 193). England: John Wiley & Sons.

Iidaka, T., Sadato, N., Yamada, H., & Yonekura, Y. (2000). Functional asymmetry of human prefrontal cortex in verbal and non-verbal episodic memory as revealed by fMRI. *Cognitive Brain Research, 9,* 73 – 83.

Ivison, D. (1977). The Wechsler Memory Scale: Preliminary findings toward an Australian standardisation. *Australian Psychologist, 12,* 303 – 313.

Ivison, D. (1990). Reliability (stability) study of the Wechsler Memory Scale, form 2. *The Clinical Neuropsychologist, 4,* 375 – 378.

Ivison, D. (1993). Towards a standardisation of the Wechsler Memory Scale form 2. *The Clinical Neuropsychologist, 7,* 268 – 280.

Ivnik, R. J., Malec, J.F., Smith, G.E., Tangalos, E. G., Petersen, R. C., Korkmen, E., & Kurland, L. T. (1992). Mayo's older Americans normative studies: WMS – R norms for ages 56 – 94. *The Clinical Neuropsychologist, 6,* 49 – 82.

Jones-Gotman, M. (1986). Right hippocampal excision impairs learning and recall of a list of abstract designs. *Neuropsychologia, 24,* 659 – 670.

Jones-Gotman, M., Zatorre, R. J., Olivier, A., Andermann, F., Cendes, F., Staunton, J., McMackin, D., Siegel, A. M., & Wieser, H. G. (1997). Learning and retention of words and designs following excision from medial or lateral temporal lobe structures. *Neuropsychologia, 35,* 963 – 973.

Kapur, N. (1988). *Memory Disorders in Clinical Practice.* London: Butterworths.

Kapur, N. (1993). Focal retrograde amnesia in neurological disease: A critical review. *Cortex, 29,* 217 – 234.

Kazui, H., Hashimoto, M., Hirono, N., & Mori, E. (2003). Nature of personal semantic memory: Evidence from Alzheimer's disease. *Neuropsychologia, 41,* 981 – 988.

Kinsbourne, M., & Wood, F. (1975). Short-term memory processes and the amnesic syndrome. In D. Deutsch & J.A. Deutsch (Eds.), *Short-term memory*. New York: Academic Press.

Klingberg, T., & Roland, P. E. (1998). Right prefrontal activation during encoding, but not during retrieval, in a non-verbal paired associates tasks. *Cerebral Cortex, 8,* 73 – 79.

Knowles, E. (2004). *Oxford Dictionary of Quotations* (6th ed.). New York: Oxford University Press.

Kopelman, M. D. (2002). Disorders of memory. *Brain, 125,* 2152 – 2190.

LaBar, K. S., & LeDoux, J. E. (2003). Emotion and the brain: An overview. In T. E. Feinberg & M. J. Farah (Eds.), *Behavioral Neurology and Neuropsychology* (2nd ed., pp. 711 – 724).

Lacritz, L. H., & Cullum, C. M. (2003). The WAIS – III and WMS – III: Practical issues and frequently asked questions. In D. S. Tulsky., D. H. Saklofske., G. J. Chelune., R. K. Heaton., R. J. Ivnik., R. Bornstein., A. Prifitera., & M. F. Ledbetter (Eds.), *Clinical Interpretation of the WAIS – III and WMS – III* (pp. 491 – 532). USA: Academic Press.

Larrabee, G. J. (1999). Book and test review: The WMS-III. *Archives of Clinical Neuropsychology, 14,* 473-477.

Larrabee, G. J. (2000). Association between IQ and neuropsychological test performance: Commentary on Tremont, Hoffman, Scott, and Adams. (1998). *The Clinical Neuropsychologist, 14,* 139-145.

Larrabee, G. J., & Curtiss, G. (1995). Construct validity of various verbal and visual memory tests. *Journal of Clinical and Experimental Neuropsychology, 17,* 536 – 547.

Larrabee, G. J., Kane, R. L., Schuck, J. R., & Francis, D. E. (1985). The construct validity of various memory testing procedures. *Journal of Clinical and Experimental Neuropsychology, 7,* 239 – 250.

LeDoux, J. E. (1992). Emotion as memory: Anatomical systems underlying indelible neural traces. In S.-A. Christianson (Ed.), *The Handbook of Emotion and Memory: Research and Theory* (pp. 269 – 288). Hillsdale, NJ: Erlbaum.

Lee, A. C. H., Robbins, T. W., Pickard, J. D., & Owen, A. M. (2000). Asymetric frontal activation during episodic memory: The effects of stimulus type on encoding and retrieval. *Neuropsychologia, 38,* 677 – 692.

Lee, T. M. C., Yip, J. T. H., & Jones-Gotman, M. (2002). Memory deficits after resection from left or right anterior temporal lobe in humans: A meta-analytic review. *Epilepsia, 43*, 283 – 291.

Leonberger, F. T., Nicks, S. D., Larrabee, G. J., & Goldfader, P. R. (1992). Factor structure of the Wechsler Memory Scale –Revised within a comprehensive neuropsychological battery. *Neuropsychology, 6*, 239 – 249.

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

Lezak, M. D., Howieson, D. B., & Loring, D. W. (2004). *Neuropsychological Assessment* (4th ed.). New York: Oxford University Press.

Lineweaver, T. T., & Chelune, G. J. (2003). Use of the WAIS-III and WMS-III in the context of serial assessments: Interpreting reliable and meaningful change. In D. S. Tulsky., D. H. Saklofske., G. J. Chelune., R. K. Heaton., R. J. Ivnik., R. Bornstein., A. Prifitera., & M. F. Ledbetter (Eds.), *Clinical Interpretation of the WAIS – III and WMS – III* (pp. 303 – 337). USA: Academic Press.

Little, T. D., Cunningham, W.A., Shahar, G., & Widaman, K.F. (2002). To parcel or not to parcel: Exploring the question, weighting the merits. *Structural Equation Modeling, 9*(2), 151-173.

Loring, D. W. (1989). The Wechsler Memory Scale – Revised, or the Wechsler Memory Scale – Revisited? *The Clinical Neuropsychologist, 3*, 59 – 69.

Loring, D. W., Lee, G. P., Martin, R. C., & Meador, K. J. (1989). Verbal and visual memory index discrepancies from the Wechsler Memory Scale – Revised: Cautions in interpretation. *Psychological Assessment, 1*, 198 – 202.

Loring, D.W., & Papanicolaou, A. C. (1987). Memory assessment in neuropsychology: theoretical considerations and practical utility. *Journal of Clinical and Experimental Neuropsychology, 9*, 340 – 358.

Markowitsch, H. J. (2000). Neuroanatomy of memory. In E. Tulving., & Craik, F. I. M (Eds.), *The Oxford Handbook of Memory* (pp. 465 – 484). New York: Oxford University Press.

Markowitsch, H. J. (2003). Functional neuroanatomy of learning and memory. In P. W. Halligan., U. Kischka., & J. C. Marshall (Eds.), *Handbook of Clinical Neuropsychology* (pp. 724 – 740) New York: Oxford University Press.

Mayes, A. R. (1995). The assessment of memory disorders. In A. D. Baddeley, B. A. Wilson, & F. N. Watts (Eds.), *Handbook of Memory Disorders* (pp. 367 – 392). England: John Wiley & Sons.

Mayes, A. R. (1986). Learning and memory disorders and their assessment. *Neuropsychologia, 24*, 25 – 39.

Mayes, A. R., & Montaldi, D. (2001). Exploring the neural bases of episodic and semantic memory: the role of structural and functional neuroimaging. *Neuroscience and Behavioral Review, 25*, 555 – 573.

Mayes, A. R., & Montaldi, D. (1999). The neuroimaging of long-term memory encoding processes. *Memory, 7*, 613 – 660.

McDaniel, M. A., & Einstein, G. O. (2007). *Prospective memory: An overview and synthesis of an emerging field.* USA: Sage Publications.

McGrew, K.S. (1997). Analysis of the major intelligence batteries according to a proposed comprehensive Gf-Gc framework. In Flanagan, D.P., Genshaft, J.L., & Harrison, P. (Eds.) *Contemporary intellectual assessment: Theories, tests and issues.* pp. 151 – 179. New York: Guilford Press.

McGuire, B. E., & Batchelor, J. (1998). Inter-rater reliability of the WMS – R Logical Memory and Visual Reproduction subtests in a neurosurgical sample. *Australian Psychologist, 33*, 231 – 233.

Meyers, J. E., & Meyers, K. R. (1995). *Rey Complex Figure Test and Recognition Trial Professional Manual.* Odessa: Psychological Assessment Resources Inc.

Milberg, W. P., Hebben, N., Kaplan, E. (1986). The Boston process approach to neuropsychological assessment. In I. Grant & K. M. Adams (Eds.), *Neuropsychological Assessment of Neuropsychiatric Disorders* (pp. 56 – 86). New York: Oxford University Press.

Miller, L. A., Lai, R., & Munoz, D. G. (1998). Contributions of the entorhinal cortex, amygdala and hippocampus to human memory. *Neuropsychologia, 36,* 1247 – 1256.

Millis, S. R., Malina, A. C., Bowers, D. A., & Ricker, J. H. (1999). Confirmatory factor analysis of the Wechsler Memory Scale – III. *Journal of Clinical and Experimental Neuropsychology, 21*, 87 – 93.

Milner, B. (1971). Interhemispheric differences in the localization of psychological processes in man. *British Medical Bulletin, 27,* 272 – 277.

Mishkin, M. (1982). *A memory system in the monkey*. Philosophical Transactions of the Royal Society of London, B298, 85 – 92.

Mishkin, M. (1978). Memory in monkeys severely impaired by combined but not by separate removal of amygdala and hippocampal formation. *Nature*, *273,* 297 – 298.

Mishkin, M., Vargha-Khadem, F., & Gadian D. G. (1998). Amnesia and the organization of the hippocampal system. *Hippocampus, 8,* 212 – 216.

Mitchell, M. (1987). Scoring discrepancies on two subtests of the Wechsler Memory Scale. *Journal of Consulting and Clinical Psychology, 55*, 914 – 915.

Mittenberg, W., & Burton, D.B. (1992). Normative data for the Wechsler Memory Scale – Revised: 25 to 34 year olds. *Psychological Assessment, 4*, 363 – 368.

Mitrushina, M. N., Boone, K. B., & D'Elia, L. F. (1999). *Handbook of Normative Data for Neuropsychological Assessment.* New York: Oxford University Press.

Mitrushina, M. N., Boone, K. B., Razani, J., & D'Elia, L. F. (2005). *Handbook of Normative Data for Neuropsychological Assessment* (2nd ed.). New York: Oxford University Press.

Miyake, A., Friedman, N. P., Shah, P., Rettinger, D. A., & Hegarty, M. (2001). How are visuospatial working memory, executive functioning, and spatial abilities related? A latent-variable analysis. *Journal of Experimental Psychology, 20,* 621 – 641.

Moye, J. (1997). Nonverbal memory assessment with designs: Construct validity and clinical utility. *Neuropsychology Review, 7,* 157 – 170.

Mummery, C. J., Patterson, K., Wise, R. J. S., Price, C. J., Hodges, J. R., & Vandenbergh, R. (1999). Disrupted temporal lobe connections in semantic dementia. *Brain, 122*, 61 – 73.

Nadel, L., & Jacobs, W. J. (1998). Traumatic memory is special. *Current Directions in Psychological Science, 7*, 154 – 157.

Nadel, L. (1994). Multiple memory systems: What and why, an update. In D. L. Schacter & E. Tulving (Eds.), *Memory Systems* (pp. 39 – 58). Cambridge, MA: MIT Press.

Nadel, L. (1992). Multiple memory systems: What and why. *Journal of Cognitive Neuroscience, 4*, 179 – 188.

Nadel, L., & Moscovitch, M. (1997). Memory consolidation and the hippocampal complex. *Current Opinion in Neurobiology, 7*, 217 – 227.

Naugle, R. I., Chelune, G. J., Cheek, R., Luders, H., & Awad, I. A. (1993). Detection of changes in material – specific memory following temporal lobectomy using the Wechsler Memory Scale – Revised. *Archives of Clinical Neuropsychology, 8*, 381 – 395.

Nott, P. N. (1975). The paired-associate learning subtest of the Wechsler Memory Scale: Six new parallel forms. *British Journal of Social and Clinical Psychology, 14*, 199 – 201.

Nunn, J. A., Polkey, C. E., & Morris, R. G. (1998). Selective spatial memory impairment after right unilateral temporal lobectomy. *Neuropsychologia, 36*, 837 – 848.

Nyberg, L., McIntosh, A. R., & Tulving, E. (1998). Functional brain imaging of episodic and semantic memory with positron emission tomography. *Journal of Molecular Medicine, 76*, 48 – 52.

O'Connor, M., & Verfaellie, M. (2004). The amnesic syndrome: Overview and Subtypes. In A. D. Baddeley., M. D. Kopelman., & B. A. Wilson (Eds.), *The Essential Handbook of Memory Disorders* (pp. 15 – 36). England: John Wiley & Sons.

Osterlind, S. J. (2006). *Modern Measurement: Theory, principles, and applications of mental appraisal.* New Jersey, US: Pearson Education.

Papanicolaou, A. C. (2006). *The amnesias: A clinical textbook of memory disorders.* USA: Oxford University Press.

Parkin, A. J. (1982). Residual learning capability in organic amnesia. *Cortex, 18,* 417 – 430.

Parkin, A. J. (1997). *Memory and Amnesia: An introduction* (2nd ed.). Oxford: Blackwell Publishers Ltd.

Parkin, A. J., & Leng, N. R. C. (1993). *Neuropsychology of the Amnesic Syndrome.* East Sussex: Lawrence Erlbaum Associates.

Penfield, W., & Perot, P. (1963). The brains record of auditory and visual experience. *Brain, 86,* 595 – 696.

Petersen, L. R., & Petersen, M. J. (1959). Short term retention of individual items. *Journal of Experimental Psychology, 91,* 341 – 343.

Price, L. R., Said, K., & Haaland, K. Y. (2004). Age-associated memory impairment of Logical Memory and Visual Reproduction. *Journal of Clinical and Experimental Neuropsychology, 26,* 531 – 538.

Price, L. R., Tulsky, D., Millis, S., & Weiss, L. (2002). Redefining the factor structure of the Wechsler Memory Scale – III: Confirmatory factor analysis with cross-validation. *Journal of Clinical and Experimental Neuropsychology, 24,* 574 – 585.

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

Reber, P. J., Knowlton, B. J., & Squire, L. R. (1996). Dissociable properties of memory systems: Differences in the flexibility of declarative and non-declarative knowledge. *Behavioural Neuroscience, 110*, 861 – 871.

Redoblado, M. A., Grayson, S. J., & Miller, L. A. (2003). Lateralised-temporal-lobe-Lesion effects on learning and memory: Examining the contributions of stimulus novelty and presentation mode. *Journal of Clinical and Experimental Neuropsychology, 25,* 36 – 48.

Rempel-Clower, N. L., Zola, S. M., Squire, L. R., & Amaral, D. G. (1996). Three cases of enduring memory impairment after bilateral damage limited to the hippocampal formation. *Journal of Neuroscience, 16,* 5233 – 5255.

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

Roid, G. H., Prifitera, A., & Ledbetter, M. (1988). Confirmatory analysis of the factor structure of the Wechsler Memory Scale – Revised. *The Clinical Neuropsychologist, 2*, 116 – 120.

Rosenbaum, S. R., Kohler, S., Schacter, D. L., Moscovitch, M., Westmacott, R., Black, S. E., Gao, F., & Tulving, E. (2005). The case of K.C.: Contributions of a memory impaired person to memory theory. *Neuropsychologia, 43*, 989 – 1021.

Rotenberg, V. S., Weinberg, I. (1999). Human memory, cerebral hemispheres, and the limbic system: A new approach. *Genetic, Social & General Psychology Monographs, 125,* 45 – 71.

Roth, D. L., Conboy, T. J., Reeder, K. P., & Boll, T. J. (1990). Confirmatory factor analysis of the Wechsler Memory Scale –Revised in a sample of head injured patients. *Journal of Clinical and Experimental Neuropsychology, 12,* 834 – 842.

Rugg, M. D., Fletcher, P. C., Frith, D. C., Frackowiak, R. S. J., & Dolan, R. J. (1996). Differential activation of the prefrontal cortex in successful and unsuccessful memory retrieval. *Brain, 119*, 2073 – 2083.

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

Schacter, D. L., Chiu, P. C-Y., & Ochsner, K. N. (1993). Implicit memory: A selective review. *Annual Review Neuroscience, 16*, 159 – 182.

Schacter, D. L., & Tulving, E. (1994). What are the memory systems of 1994? In D. L. Schacter & E. Tulving (Eds.), *Memory Systems* (pp. 1 – 38). Cambridge, MA: MIT Press.

Schmidt, F. L., & Hunter, J. E. (1996). Measurement error in psychological research: Lessons from 26 research scenarios. *Psychological Methods, 1,* 199 – 223.

Scoville, W. B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of Neurology, Neurosurgery, Psychiatry, 20,* 11 – 21.

Shallice, T., and Warrington, E. K. (1970). Independent functioning of verbal memory stories: A neuropsychological study. *Quarterly Journal of Experimental Psychology*, *22*, 261 – 273.

Shores, A. E., & Carstairs, J. R. (2000). The Macquarie University neuropsychological normative study (MUNNS): Australian norms for the WAIS – R and WMS – R. *Australian Psychologist, 35*, 41 – 59.

Skilbeck, C. E., & Woods, R. T. (1980). The factorial structure of the Wechsler Memory Scale: Samples of neurological and psychogeriatric patients. *Journal of Clinical Neuropsychology, 2,* 293 – 300.

Smith, G. E., Ivnik, R. J., Malec, J. F., Petersen, R. C., Tangalos, E. G., & Kurland, L. T. (1992). Mayo's Older American Normative Studies (MOANS): Factor structure of a core battery. *Psychological Assessment, 4,* 382-390.

Smith, G. E., Malec, J. F., & Ivnik, R. J. (1992). Validity of the construct of nonverbal memory: A factor analytic study in a normal elderly sample. *Journal of Clinical and Experimental Neuropsychology*, *14,* 211 – 221.

Snowden, J. (2002). Disorders of semantic memory. In A. D. Baddeley, M. D., Kopelman & B. A. Wilson (Eds.), *The Handbook of Memory Disorders* (2nd ed.). (pp. 293 – 314). Chichester, England: Wiley.

Spreen, O., & Strauss, E. (1998). *A Compendium of Neuropsychological Tests* (2nd ed.). New York: Oxford University Press.

Squire, L. R. (1986). The neuropsychology of memory dysfunction and its assessment. In I. Grant., & K. M. Adams (Eds.), *Neuropsychological Assessment of Neuropsychiatric Disorders* (pp. 245 – 312). New York: Oxford University Press.

Squire, L. R. (1987). *Memory and Brain.* New York: Oxford.

Squire, L. R. (1992). Declarative and non-declarative memory: Multiple brain systems supporting learning and memory. *Journal of Cognitive Neuroscience, 4,* 232 – 243.

Squire, L. R., Amaral, D. G., & Press, G. A. (1990). Magnetic resonance imaging of hippocampal formation and mammillary nuclei distinguish medial temporal lobe and diencephalic amnesia. *Journal of Neuroscience, 10,* 3106 – 3117.

Squire, L. R., & Butters, N. (1992). *Neuropsychology of Memory* (2nd ed.). New York: Guilford Press.

Squire, L. R., Knowlton, B., & Musen, G. (1993). The structure and organization of memory. *Annual Review Psychology, 44,* 453 – 495.

Squire, L. R., & Shimamura, A. (1996). The neuropsychology of memory dysfunction and its assessment. In I. Grant., & K. M.Adams (Eds.), *Neuropsychological Assessment of Neuropsychiatric Disorders* (pp. 232 – 262). New York: Oxford University Press.

Squire, L. R., & Zola, S. M. (1998). Episodic memory, semantic memory and amnesia. *Hippocampus, 8,* 205 – 211.

Tabachnick, B. G., & Fidell, L. S. (1996). *Using Multivariate Statistics* (3rd ed.). New York: Harper Collins Publishers.

Taylor, E. M. (1959). *Psychological Appraisal of Children with Cerebral Deficits.* Cambridge, MA: Harvard University Press.

Thorndike, R. L., Hagen, E. P., & Sattler, J. M. (1986). *Stanford-Binet Intelligence Scale* (4th ed.). Chicago, IL: Riverside Publishing Co.

Troster, A. I., Butters, N., Salmon, D. P., Cullum, C. M., Jacobs, D., Brandt, J. & White, R. F. (1993). The diagnostic utility of saving scores: Differentiating Alzheimer's & Huntington's diseases with the logical memory & visual reproduction tests. *Journal of Clinical & Experimental Neuropsychology, 15,* 773 – 788.

Tulsky, D.S. (2004). A new look at the WMS – III: New research to guide clinical practice. *Journal of Clinical and Experimental Neuropsychology, 26,* 453 – 458.

Tulsky, D. S., Chelune, G. J., & Price, L. R. (2004). Development of a new delayed Memory Index for the WMS – III. *Journal of Clinical and Experimental Neuropsychology, 26*, 563 – 576.

Tulsky, D. S., Chiaravalloti, N. D., Palmer, B. W., & Chelune, G. J. (2003). The Wechsler Memory Scale, Third Edition: A new perspective. In D. S. Tulsky., D. H. Saklofske., G. J. Chelune., R. K. Heaton., R. J. Ivnik., R. Bornstein., A. Prifitera., & M. F. Ledbetter (Eds.), *Clinical Interpretation of the WAIS – III and WMS – III* (pp. 93 – 139). USA: Academic Press.

Tulsky, D.S., Ivnik, R.J., Price, L.R., & Wilkins, C. (2003). Assessment of cognitive functioning with the WAIS-III and WMS-III: Development of a six factor model. In D.S. Tulsky, D.H. Saklofske, R. Bornstein, G. Chelune, R. Heaton, R. Ivnik, M. Ledbetter, & A. Prifitera. *Clinical Interpretation of the WAIS-III and WMS-III* (pp. 147-179). San Diego: Academic Press.

Tulsky, D. S., & Ledbetter, M. F. (2000). Updating to the WAIS – III and WMS – III: Considerations for research and clinical practice. *Psychological Assessment, 12*, 253 – 262.

Tulsky, D.S., & Price, L.R. (2003). The joint WAIS-III and WMS-III factor structure: Development and cross-validation of a six factor model of cognitive functioning. *Psychological Assessment, 15,* 149-162.

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving and W. Donaldson (Eds.), *Organization of Memory*. New York: Academic Press.

Tulving, E. (1983). *Elements of Episodic Memory*. Clarendon: Oxford University Press.

Tulving, E. (1985). How many memory systems are there? *American Psychologist, 40,* 385 – 398.

Tulving, E. (2002). Episodic memory: from mind to brain. *Annual Reviews in Psychology, 53,* 1 – 25.

Tulving, E., & Markowitsch, H. J. (1998). Episodic and declarative memory: Role of the hippocampus. *Hippocampus, 8,* 198 – 204.

Tulving, E., & Schacter, D.L. (1990). Priming and human memory systems. *Science, 247,* 301 – 307.

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80,* 352 – 373.

Van der Werf, Y. D., Scheltens, P., Lindeboom, J., Witter, M. P., Uylings, H. B. M., & Jolles, J. (2003). Deficits of memory, executive functioning and attention following infarction in the thalamus: A study of 22 cases with localized lesions. *Neuropsychologia, 41,* 1330 – 1344.

Van der Werf, Y. D., Witter, M., Uylings, H., & Jolles, J. (2000). Neuropsychology of infarctions in the thalamus: A review. *Neuropsychologia, 38,* 613 – 627.

Victor, M., Adams, R. D., & Collins, G. H. (1989). *The Wernicke-Korsakoff Syndrome and Related Neurologic Disorders due to Alcoholism and Malnutrition* (2nd ed.). Philadelphia, PA: Davis.

Waldmann, B. W., & Dickson, A. L. (1991). The relationship between intellectual function and performance on the Wechsler Memory Scale. *Journal of Genetic Psychology, 152,* 57 – 70.

Walsh, K., & Darby, D. (2001). *Neuropsychology: A clinical approach* (4th ed.). England: Harcourt Publishers Ltd.

Warrington, E. K. (1984). *Recognition Memory Test Manual.* Windsor, England: NFER-Nelson.

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

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

Wechsler, D. (1981). *Wechsler Adult Intelligence Scale – Revised.* New York: Psychological Corporation.

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

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale – Third Edition*. San Antonio, TX: The Psychological Corporation.

Wechsler, D. (1997). *Wechsler Memory Scale – Third Edition.* San Antonio, TX: The Psychological Corporation.

Wiggs, C. L., Weisberg, J., & Martin, A. (1999). Neural correlates of semantic and episodic memory retrieval. *Neuropsychologia, 37,* 103 – 118.

Wilde, N. J., Strauss, E., Chelune, G. J., Hermann, B. P., Hunter, M., Loring, D.W., Martin, R. C., & Sherman, E. M. S. (2003). Confirmatory factor analysis of the WMS – III in patients with Temporal Lobe Epilepsy. *Psychological Assessment, 15,* 56 – 63.

Williams, M. A., Rich, M. A., Reed, L. K., Jackson, W. T., LaMarche, J. A., & Boll, T. J. (1998). Visual Reproduction subtest of the Wechsler Memory Scale-Revised: Analysis of construct validity. *Journal of Clinical Psychology, 54,* 963 – 971.

Willmes, K. (2003). The methodological and statistical foundations of neuropsychological assessment. In P. W. Halligan., U. Kischka., & J. C. Marshall (Eds.), *Handbook of Clinical Neuropsychology* (pp. 27 – 47). New York: Oxford University Press.

Williams, J. M. (1991). *Memory Assessment Scales manual.* Odessa: Psychology Assessment Resources.

Wilson, B. A. (2004). Assessment of memory disorders. In A. D. Baddeley., M. D. Kopelman., & B. A. Wilson (Eds.), *The Essential Handbook of Memory Disorders* (pp. 159 – 178). England: John Wiley & Sons.

Winocur, G., Moscovitch, M., Caruana, D. A., & Binns, M. A. (2005). Retrograde amnesia in rats with lesions to the hippocampus on a test of spatial memory. *Neuropsychologia, 43,* 1580 – 1590.

Wirsen, A., & Ingvar, D. H. (1991). Memory deficits in patients with chronic frontal lesions. *Neuropsychology, 5,* 151 – 169.

Wong, J. L., & Gilpin, A. R. (1993). Verbal vs visual categories on the Wechsler Memory Scale – Revised: How meaningful a distinction? *Journal of Clinical Psychology, 49,* 847 – 854.

Wood, F., Ebert, V., & Kinsbourne, M. (1982). The episodic-semantic memory distinction in memory and amnesia: clinical and experimental observations. In L. S. Cermak (Ed.), *Human Memory and Amnesia* (pp. 167 – 194). New Jersey: Erlbaum.

Woodard, J. L., & Axelrod, B. N. (1995). Parsimonious prediction of Wechsler Memory Scale – Revised memory indices. *Psychological Assessment, 7,* 445 – 449.

Zhu, J., Tulsky, D. S. (2000). Co-norming the WAIS – III and WMS – III: Is there a test-order effect on IQ and memory scores? *The Clinical Neuropsychologist, 14,* 461 – 467.

Zola, M. S. (1998). Memory, amnesia, and issue of recovered memory: Neurobiological aspects. *Clinical Psychology Review, 18,* 915 – 932.

Zola-Morgan, S., & Squire, L. R. (1985). Medial temporal lesions in monkeys impair memory on a variety of tasks sensitive to human amnesia. *Behavioural Neuroscience, 99,* 22 – 34.

Zola-Morgan, S., & Squire, L. R. (1990). The primate hippocampal formation: Evidence for a time-limited role in memory storage. *Science, 250,* 288 – 291.

Zola-Morgan, S., & Squire, L. R. (1991). The medial temporal lobe memory system. *Science, 253,* 1380 – 1387.

Zola-Morgan, S., & Squire, L. R. (1993). Neuroanatomy of memory. *Annual Review of Neuroscience*, *16*, 547 – 563.

Zola-Morgan, S., Squire, L. R. & Amaral, D. (1986). Human amnesia and the medial temporal region: Enduring memory impairment following a bilateral lesion limited to field CA1 of the hippocampus. *Journal of Neuroscience, 6,* 2950 – 2 967.

Zola-Morgan, S., Squire, L. R., Amaral, D. G., & Suzuki, W. A. (1989). Lesions of perirhinal and parahippocampal cortex that spare the amygdale and hippocampal formation produce severe memory impairment. *Journal of Neuroscience, 9,* 4355 – 4370.

Appendix A

The Original Scoring System of the Visual Reproduction subtest of the Wechsler Memory Scale – Revised

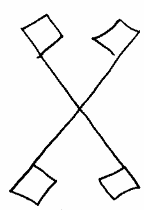
Score ranges, scoring criteria, and sample drawings for each Visual Reproduction card are presented below. Generally, one point is awarded for each criterion met.

A sample drawing that would receive full credit is presented for each card, followed by an explanation of each criterion and several sample drawings that would *not* receive credit for that criterion. Nearly all of the drawings presented were taken from actual standardization records.

Unless otherwise noted, the criteria are each applied independently to the drawings. For example, a drawing for Card A need not receive credit for Criterion 3 in order to receive credit for Criterion 4.

Some of the criteria specify numerical limits or measurements by which a drawing should be judged; however, it will not be necessary to take actual measurements in most cases. The precise numerical rules in scoring the standardization records and preparing the norms are included so that examiner's may duplicate the standardization procedures when scoring a particularly difficult drawing.

Card A (Score Range: 0-7)

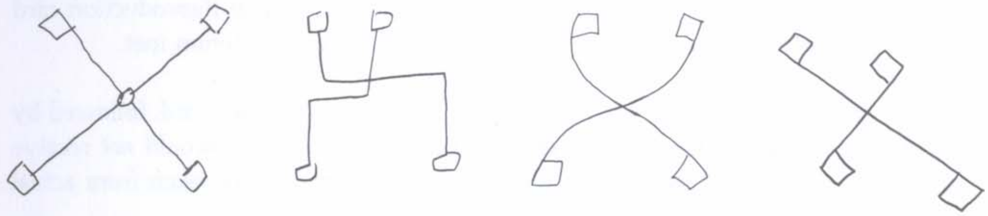


Sample full-credit drawing

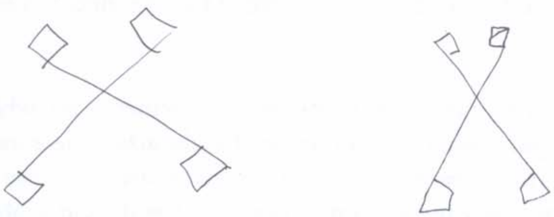
Scoring Criteria:

Staffs (a single staff, and staffs that do not cross, are not eligible for scoring and receive no credit)

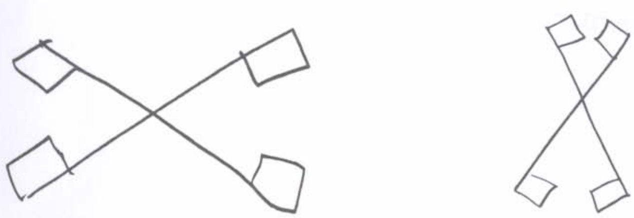
1. Lines forming the two staffs are unbroken, reasonably straight, and approximately equal in length (the ratio of the length of the longer staff to the shorter must not exceed 1.5 and is to be measured from the far ends of the staffs).



2. Staffs intersect at the midpoints of each (the ratio of the longer arm to the shorter arm of each staff from its point of intersection must not exceed 1.5).



3. Staffs cross at right angles (no angle formed is less than 75°).

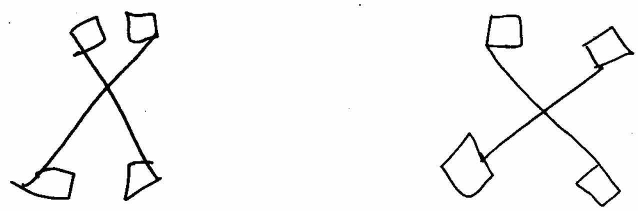


4. Figure has not been rotated more than 15°.

Flags (fewer than four are not eligible for scoring and receive no credit for Criteria 5-7)

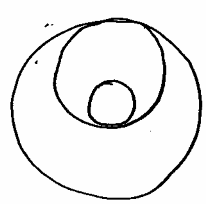
5. All four flags point in the correct direction.

6. All four flags share a side with a staff.



7. All four flags are square in shape (a. the ratio between the longest and shortest side must not exceed 1.5; b. there must be four right angles with none less than 75°).

Card B (Score Range: 0-7)

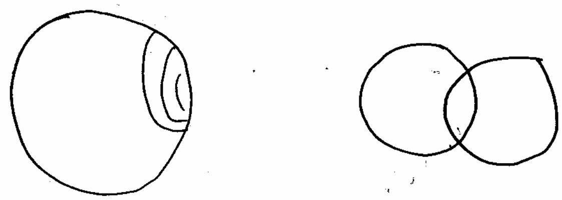


Sample full-credit drawing

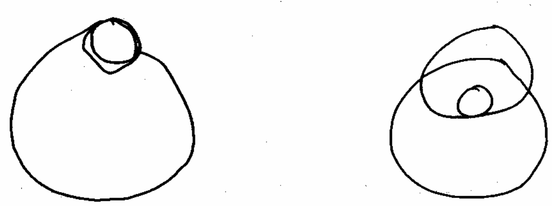
Scoring Criteria

Circles (a single circle is not eligible for scoring and receives a score of 0)

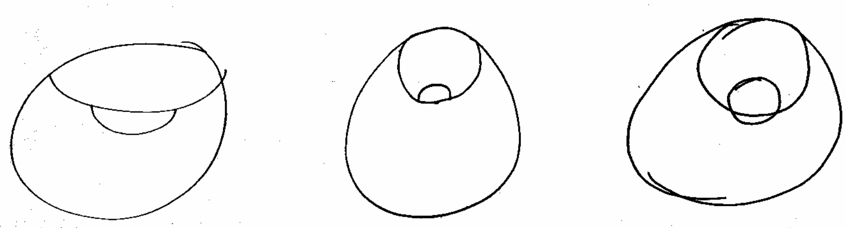
1. A large circular figure that contains at least one other circular figure.



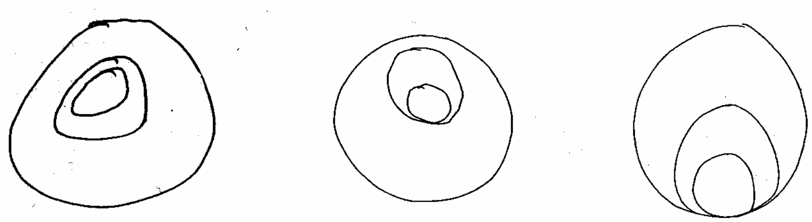
2. A middle-sized circular figure is inside the large circular figure.



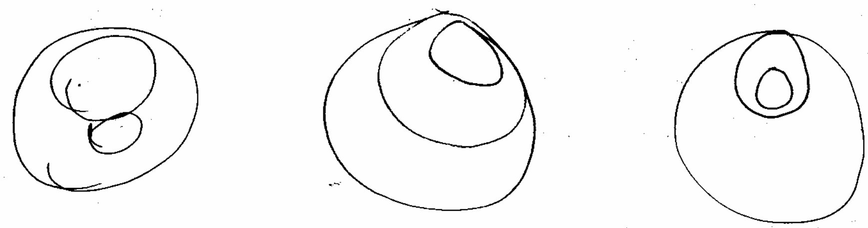
3. A small circular figure is inside the middle-sized circular figure.



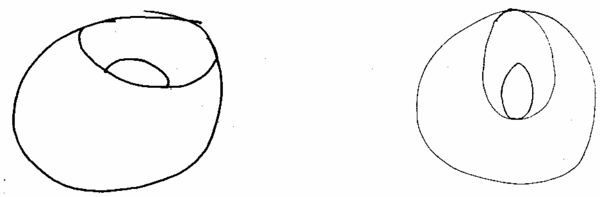
4. Middle-sized and large circular figures touch at the top of each (to be considered "touching" there must not be a measurable amount of space between the circles).



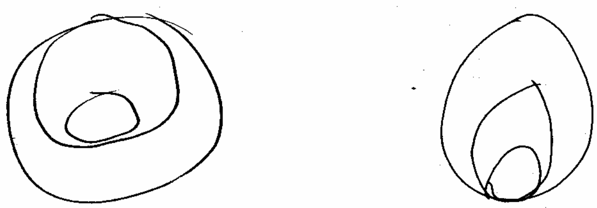
5. Small and middle-sized circular figures touch at the bottom of each (to be considered "touching" there must not be a measurable amount of space between the circles).



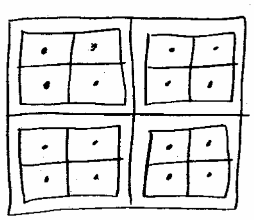
*6. All three figures are round in shape (for each figure, the longest diameter should be no more than 1.5 times the shortest diameter); all figures must be closed.



7. The relative sizes of the three figures are in correct proportion (the ratio of the longest diameter of a circle to the longest diameter of the next smaller circle is between 1.5 to 2.5).



Card C (Score Range: 0-9)

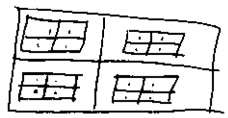
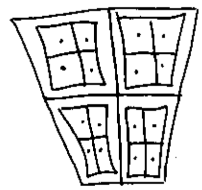


Sample full-credit drawing

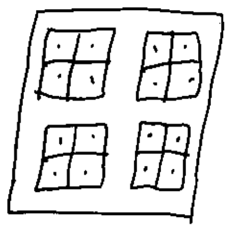
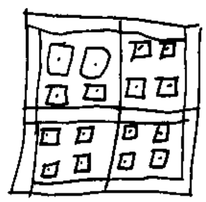
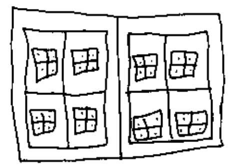
*If Criterion 6 is scored 0, Criterion 7 is also scored 0.

Large Square

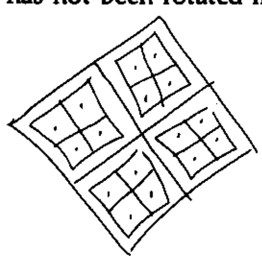
1. Large figure is square in shape (a. the ratio between the longest and shortest side must not exceed 1.5; b. there must be four right angles with none less than 75°).

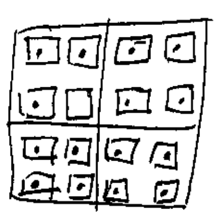
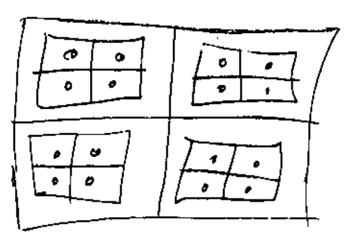
2. Large figure is bisected by both a vertical and a horizontal line, which intersect.

3. Large figure has not been rotated more than 15°.

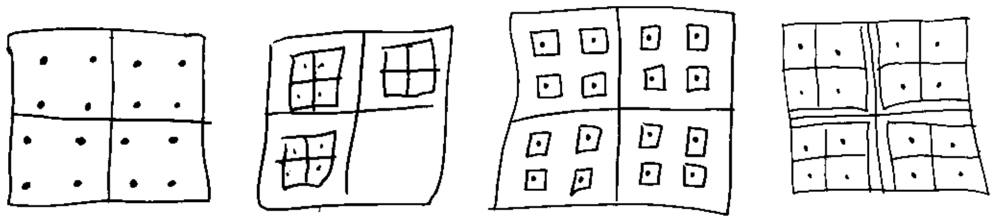


4. Each quadrant of the large figure has four dots in a square array. Vertical and horizontal diameters of the four quadrants need not be present. No dot may be a circle.

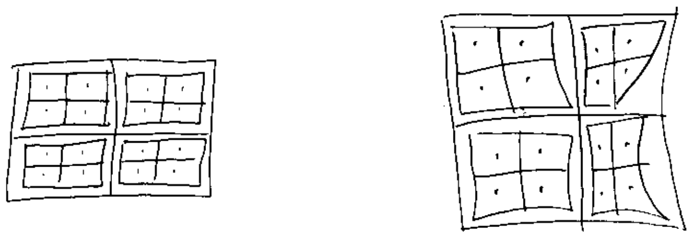



Medium-Sized Squares (fewer than 4 are not eligible for scoring)

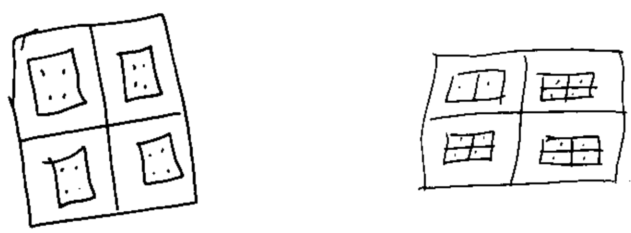
*5. One medium-sized, square-like figure is drawn in each quadrant of the large square; the medium-sized figures do not touch one another or the sides of the large square (some degree of carelessness in drawing is permissible).



6. Each of the medium figures is square in shape (a. the ratio between the longest and shortest side must not exceed 1.5; b. there must be four right angles with none less than 75°).

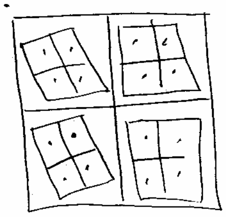


7. Each medium-sized square is bisected by both a vertical and a horizontal line, which intersect.

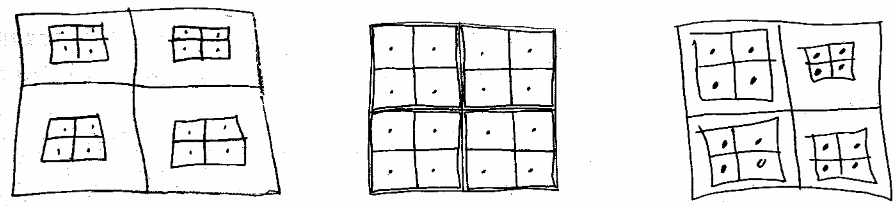


*If Criterion 5 is scored 0, Criteria 6 through 9 are all scored 0.

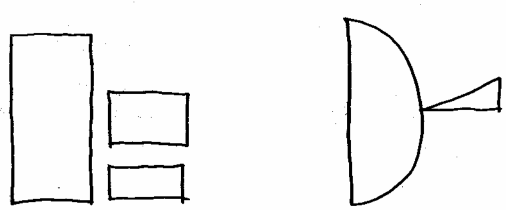
8. No medium-sized figure has been rotated more than 15°.



9. The four medium-sized squares are equal in size and are in correct proportion to the large square (the ratio of the longest diagonal of the large square to the longest diagonal of the smallest medium-sized square is between 2.5 and 3.5).



Card D (Score Range: 0-18)

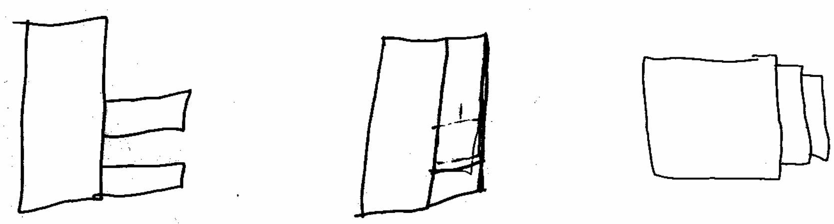


Sample full-credit drawing

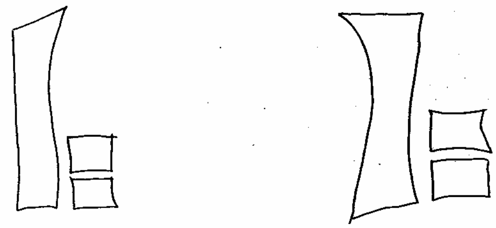
Scoring Criteria

Rectangles (a single rectangle is not eligible for scoring and receives no credit for Criteria 1-10); two rectangles are eligible to receive credit for Criteria 1-3 only).

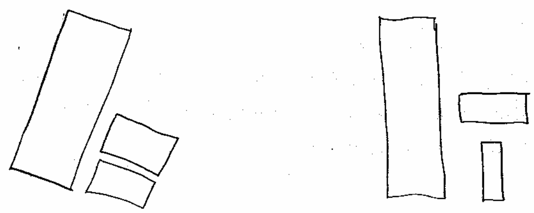
1. Rectangular figures do not touch, intersect, or contain one another.



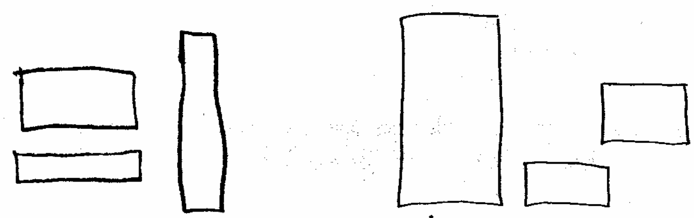
2. Interior angles are about 90° (with none less than 75°).



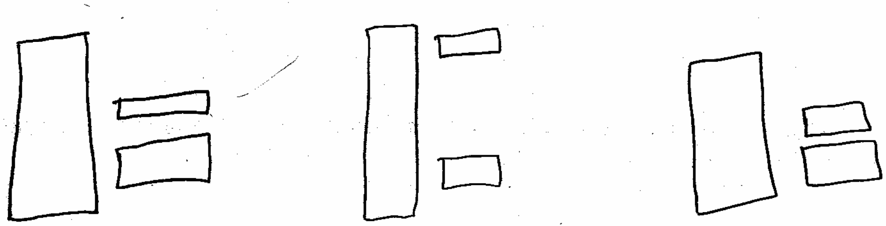
3. No figure has been rotated more than 15°.



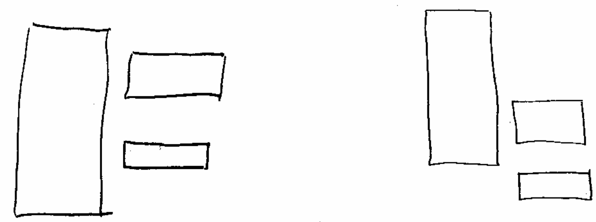
4. The two small rectangles are positioned to the right of the large rectangle, and one is directly over the other.



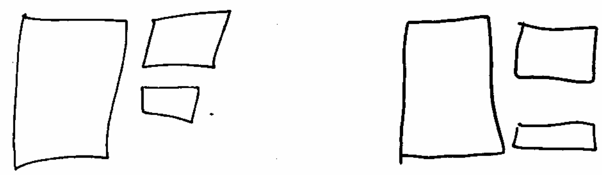
5. The upper of the two small rectangles is greater in height than the lower rectangle.



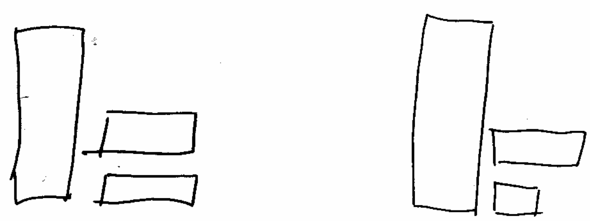
6. The base of the large rectangle is level with the base of the lower of the two small rectangles (using the longest side of the large rectangle as the reference, the base of the lower small rectangle should be less than 10% above or below the base of the large rectangle).



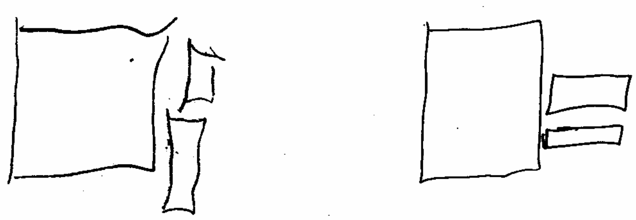
7. The top of the large rectangle is above the tops of the two small rectangles (using the longest side of the large rectangle as the reference, the top side of the upper small rectangle should be at least 10% below the top of the large rectangle).



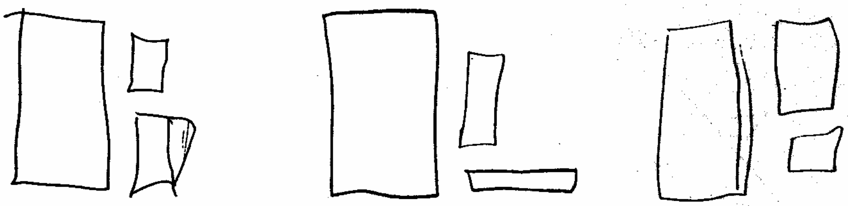
8. The bases of all three rectangles are approximately equal in length (the ratio of the longest to the shortest base must not exceed 1.5).



9. The height of the large rectangle is greater than its width (the ratio of the height to the width is between 1.5 and 4.5).



10. The heights of the two small rectangles are clearly less than their widths.

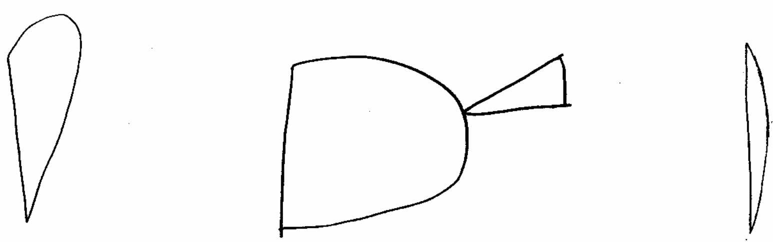


*Circle Segment

11. A figure is positioned to the right of the rectangles. (Note: *Any* figure to the right of the rectangles receives credit for this rule.)

12. The arc of the circle segment curves to the right of the vertical line.

13. The figure is reasonably symmetrical and in correct proportion.



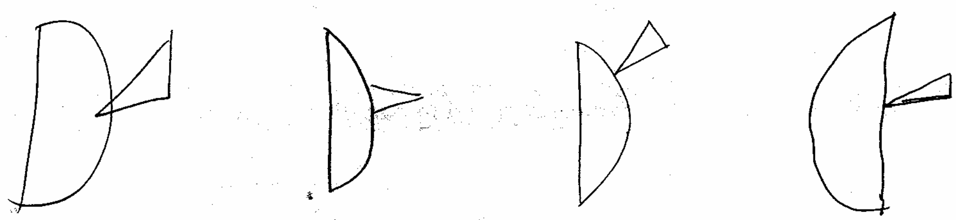
*If the figure does not resemble a circle segment (e.g., lacks a vertical line and an arc), Criteria 12 through 14, and 16, are all scored 0.

14. The figure has not been rotated more than 15°.

**Triangle

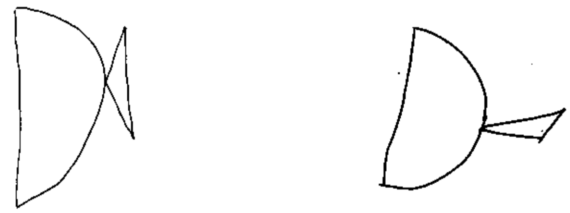
15. A figure is positioned to the right of the circle segment. (Note: *Any* figure to the right of the circle segment receives credit for this rule.)

16. A vertex of the triangle touches the midpoint of the arc (using the height of the vertical line as the reference, the point where the vertex touches the arc should be within 10% above or below an imaginary line extended horizontally from the midpoint of the vertical line).

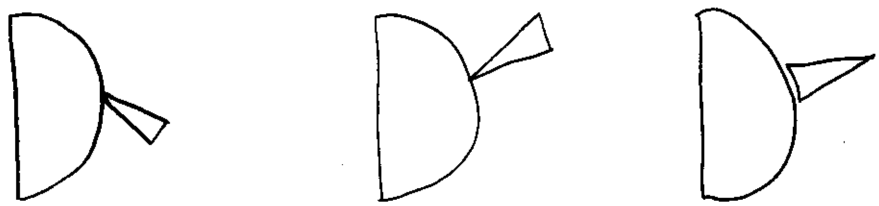


**If the figure is not a triangle (e.g., lacks 3 clear angles), Criteria 16 through 18 are all scored 0.

17. The triangle contains one angle that closely approximates 90°.



18. The triangle has not been rotated more than 15°, or revolved about one of its sides.

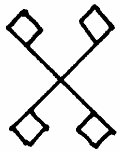


Maximum total points: 41

Appendix B

The Revised Scoring System and additional cued and recognition memory measures of the Visual Reproduction subtest of the Wechsler Memory Scale – Revised.

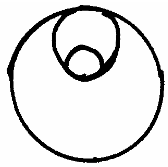
DESIGN ONE:



20 items.
Each item is worth one point.
Score each item independently.

1. There are at least two continuous lines or four lines emanating from a central point or figure. If there is only one line or more than four lines from a central point, score zero.
2. There are only two lines present and these two lines intersect. If more or less than two lines, score zero.
3. Two lines intersect in the middle 1/3 of each. A minor gap at or near the intersection is acceptable if there is no change of direction. If the figure has a radial spokes design, all spokes are of the same size (50% tolerance between longest and shortest spoke).
4. The lines that intersect or emanate from the central point do not form angles less than 45°.
5. Lines or radial spokes are similar in length. The shorter line or spoke must be at least 75% of the longest line or spoke (regardless of where they intersect).
6. The intersecting lines have not been rotated to an orthogonal position. If one line is vertical, the other is not horizontal and, if one is horizontal, the other is not vertical. If the lines to not intersect, score zero.
7. At least three geometric figures are present.
8. All figures have the same number of sides (or all are circular). Figures can share a side with the lines, but they cannot share a border with an external figure (e.g. a bordering square).
9. All figures are identical in shape and size to each other (90% tolerance).
10. Each figure is between 30%-50% of the length of the radial arm of the line. If the figure is not joined to the line, use the longest side of the figure as the reference comparison to the radial arm. If there are no lines, score zero.
11. At least 3/4 of the figures are squares (flags) i.e. they have four sides. Gaps between the line and square are acceptable if no greater than 25% of the length of the side of the figure where the gap is present.
12. Exactly four discrete figures are present. The figures do not share a border with an external square.
13. More than 50% of the figures touch any line forming a staff. Figures can overlap a line.
14. At least two figures are correctly positioned near the end of the appropriate line. A figure does not have to touch or overlap the adjacent line.
15. All four figures are correctly positioned near the end of the appropriate line. A figure does not have to touch or overlap the adjacent line.
16. All four figures touch one line at the endpoint of the line. Minor overlap or gaps (less than 10% of the longest side of the figure) are acceptable. If the figure is a circle, overshoot or gap must be less than 10% of the diameter of the circle.
17. The side of the figure is contiguous with the line (i.e. the figure shares a side with the line). This applies to all four figures. A minor gap in the line before it joins the side of the figure is acceptable (less than 10% of the length of the side of the figure).
18. Two figures face inwards (if rotation of the lines is less than 90°, assume direction that would maximise the score). If any figure is not contiguous with the line and 100% of that figure is not on the correct side of the adjacent line, score zero.
19. Four figures face inwards. If any figure is not contiguous with the line and 100% of that figure is <u>not</u> on the correct side of the adjacent line, score zero.
20. No extra elements. Minor overshoots of lines should not be penalised.

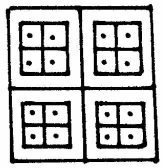
DESIGN TWO:



20 items.
Each item is worth one point.
Score each item independently.

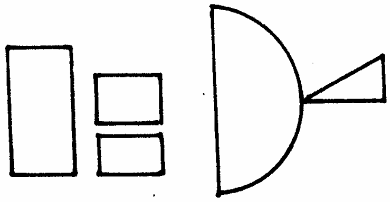
1. At least one circular figure is present.
2. Three geometric figures (only) are present and at least one is circular in shape.
3. At least two geometric figures are present, with one mostly inside the other. Figures can share a border.
4. The figures form a clear size gradient (i.e. they are not of equal size). If there are more than three figures, take the largest one to be the large figure and the smallest one to be the small figure, then choose the medium figure so as to maximise the score. If there are only two figures, they must form a clear size gradient. A dot is not a figure.

   If there are only two figures, interpret spatial relationship questions (Items 5-14) so as to maximise score.
5. Large figure mostly encloses at least one smaller figure.
6. Large figure mostly encloses two smaller figures.
7. Small figure is largely enclosed by a medium figure.
8. Medium figure is located towards the top of the large figure and away from the bottom. The gap between the bottom of the large figure and the bottom of the medium figure should be at least three times the gap between the top of those figures.
9. The top of the medium figure touches the top of the large figure. Minor overlap or gap between the figures is acceptable (less than 10% of the diameter of the large figure).
10. A small figure is located close to the bottom of the medium figure and away from the top (regardless of whether it is enclosed by the medium figure). The gap between the top of the medium figure and the top of the small figure should be at least three times the gap between the bottom of those figures.
11. The bottom of the small figure touches the bottom of the medium figure. Minor overlap or gap between the figures is acceptable (less than 10% of diameter of the medium figure).
12. Medium figure is about 1/3 the diameter of the vertical axis of the large figure. Acceptable range is 25%-50%.
13. Small figure is about 1/3 the diameter of the vertical axis of the medium figure. Acceptable range is 25%-50%.
14. Areas enclosed by each figure are in correct relative proportion. The area of the small figure is about 20%-25% of the area of the medium figure, and the area of the medium figure is about 20%-25% of the area of the large figure.
15. The figures are symmetrically placed about the midline. If a vertical midline axis is drawn to divide the largest figure, no more than 60% of any figure is on one side of that axis.
16. The spatial relationship between the three figures is preserved, even if the design is inverted. If there are only two circles, score zero.
17. All shapes are primarily closed circular figures (can be ovals). Any common circumference is less than 20% of the circumference of the larger of the two figures.
18. At least two of the figures are discrete circles rather than ovals (i.e. separate circles in their own right). For each circle, the smallest diameter is at least 90% of the largest diameter. Any common circumference is < 20% of the circumference of the larger of the two figures.
19. All figures are are discrete circles rather than ovals or any other shape. For each circle, the smallest diameter is at least 90% of the largest diameter. Any common circumference is less than 20% of the circumference of the larger of the two figures.
20. No extra elements, except minor line continuations.

DESIGN THREE: 

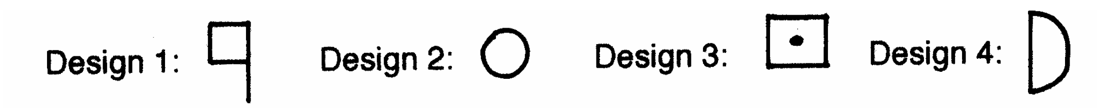
20 items.
Each item is worth one point.
Score each item independently.

1. A large figure with two or more internal elements (lines, figures) is present. The large figure may share a side with the edge of the paper for this item only. If any doubt, interpret to maximise score. If two outer squares, consider the outermost to be the large square.
2. At least one large four-sided figure is present and it is approximately square. The figure may be rectangular as long as a shorter side is at least 50% of the length of the longer sides.
3. The four sides of the large square are reasonably equal in length. The longest side is no more than 25% longer than the shortest side.
4. The four sided figure is exactly a square. Every angle is in the range 85°-95°. The smallest line is at least 90% of the largest line. Gaps or overlaps are acceptable as long as they are less than 10% of the length of the line.
5. A vertical division divides the large figure. A double-lined vertical division is acceptable. The division can be contiguous with the internal squares. Gaps in the joining of the division and the external figure are acceptable as long as the length of the vertical division is at least 75% of the vertical dimension of the square.
6. A horizontal division divides the large figure. A double-lined horizontal division is acceptable. The division can be contiguous with the internal squares. Gaps in the joining of the horizontal division and the external figure are acceptable as long as the length of the division is at least 75% of the horizontal dimension of the square.
7. The vertical and horizontal divisions intersect and divide the figure into four quadrants (i.e. they touch and cross each other).
8. Two to four smaller figures are present, with or without a major figure bordering them. Each smaller figure shares no more than two sides with any of the following: the horizontal division, the vertical division, the external square, any other smaller figure.
9. Each quadrant of the larger figure has only one smaller figure. Quadrants need not be symmetrical. No more than two figures share a line with each other. **OR** If there is no larger figure, the smaller figures form a 2x2 matrix and no more than two figures share a line.
10. Smaller figures are distinct shapes (even if there are less than or more than four figures). The figures do not overlap each other or the sides of the square or the internal divisions or any additional lines. They do not share a common border, but may touch or partly overlap.
11. Each of the smaller figures is divided into four parts, or there are four shapes in each quadrant in a 2x2 matrix. The quadrant being divided into four scores zero.
12. Each smaller figure is divided into four, or each quadrant is divided into four by a vertical and a horizontal line. Double vertical or horizontal lines are acceptable.
13. The smaller figures (as drawn by the client) are similar (90%) in configuration or size.
14. At least three of the smaller figures are in correct proportion to the larger figure, as per the original design. If there is no large square, score zero.
15. The smaller figures (as drawn by the client) have four sides and are separate from each other, from the internal divisions and the external square. There is no overlap between sides.
16. A number of dots (or circles) are present in at least 75% of the internal segments of the smaller figures, in at least 75% of each smaller figure if there are no segments, in at least 75% of the smaller figures that the client produces; or in at least 75% of the quadrants.
17. Each quadrant of the large figure has only four dots/circles in a square array. Divisions may or may not be present. If no larger figure, score zero.
18. The dot or open circle occupies less than 10% of the area in each segment of the internal figure, in each figure if no segments, or in each quadrant if no smaller figures.
19. All four smaller figures are placed symmetrically. Borders are equal and less than 20% of the length of the quadrant. If no internal divisions are present or there is no external square, equal spacing occurs between the smaller figures in both the horizontal and vertical planes.
20. No extra lines, dots or figures. Minor overshoots of lines should not be penalised.

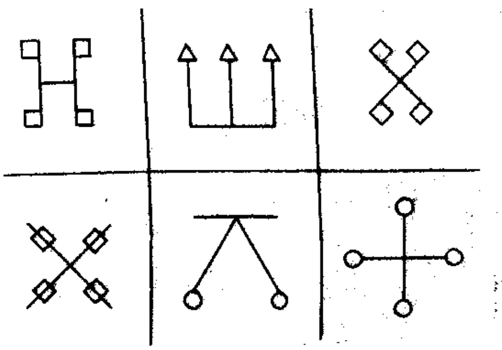
DESIGN FOUR:  20 items.
Each item is worth one point.
Score each item independently.

1. At least two figures are present, of which one is four-sided or is a circle/semicircle. Figures may share a common border. If there is only one figure, score zero.
2. A tall rectangle is present. The base of the rectangle is less than 75% of the vertical dimension. The longest side is no more than 20% longer than its parallel side.
3. One or more 3-6 sided figures is adjacent to the large rectangle (sharing a border is acceptable). **OR** There are one or more 3-6 sided figures (if no large rectangle is present).
4. The smaller figure(s) in Item 3 are separate from each other and from the major figure (rectangle). Minor touching or overlap is acceptable.
5. The bases of all figures are of similar length (the smallest base is at least 90% of the largest base). If there is only one figure, score zero.
6. The tall rectangle is clearly above the the height of the adjacent figure(s) by at least 10% of its height. The adjacent figure(s) do not need to have four sides.
7. The base of the tall rectangle and the base of the lowest adjacent figure are level (within 10% of the height of the tall rectangle). The adjacent figure(s) need not be four-sided.
8. There are two four-sided figures positioned on top of each other and to the right of the large rectangle (if it is present). The two figures' widths are greater than their heights.
9. Of the two four-sided figures in Item 8, one is clearly larger and placed above the smaller. The smaller figure is no more than 70% of the height of the figure above it, at any point.
10. A large figure with a curved surface **OR** a curved line is present.
11. The large figure in Item 10 is a discrete semicircle only (irrespective of orientation).
12. The curved portion of the semicircle faces the right.
13. The semicircle is in the correct proportion. The radius of the semicircle is half the size of the vertical dimension. (Acceptable range is 40%-60% of vertical dimension.)
14. The figure as described in Item 10 is located to the right of the figures mentioned in Items 1-9, or to the right of some other shape or even a line. If there is nothing to the left of the figure in Item 10, score zero.
15. A smaller figure is located near the figure described in Item 10. **OR** If the figure described in Item 10 is absent, a smaller figure is placed to the far right of any other shapes (as per Items 1-9). It is acceptable if the smaller figure is placed inside the figure from Item 10. A line receives credit for this item but it would not earn points on Items 16-19.
16. The smaller figure in Item 15 is separate from any other figure. The smaller figure is not inside another figure and it does not share a side with another figure. The smaller figure can touch, overlap (10% of diameter tolerance) or be in close proximity to the figure described in Item 10.
17. The smaller figure in Item 15 is located to the right of the figure described in Item 10, or to the right of some four-sided figure if the figure described in Item 10 is absent
18. The smaller figure in Item 15 is located at or near the centre of the right border of the figure in Item 10. The smaller figure must be within 30° above or below the centre of the arc of the figure in Item 10 and the smaller figure can be inside that figure. If the figure in Item 10 is absent, the smaller figure must be located above the level of the bases of the figures described in Items 1-9 and below the upper level of the edge of the large rectangle.
19. The smaller figure in Item 15 is a discrete triangle (i.e. it has three discrete sides separate from the figure described in Item 10.)
20. No extra elements. Minor overshoots of lines should not be penalised.

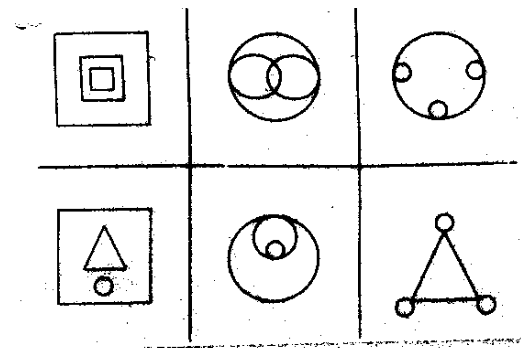
Cues for the designs of the Visual Reproduction subtest of the Wechsler Memory Scale – Revised developed by Clark (2000).



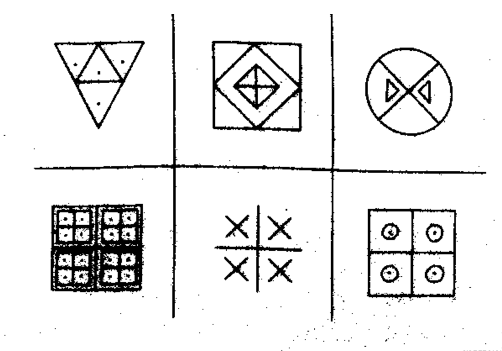
Recognition cards for the four designs of the Visual Reproduction subtest of the Wechsler Memory Scale – Revised.



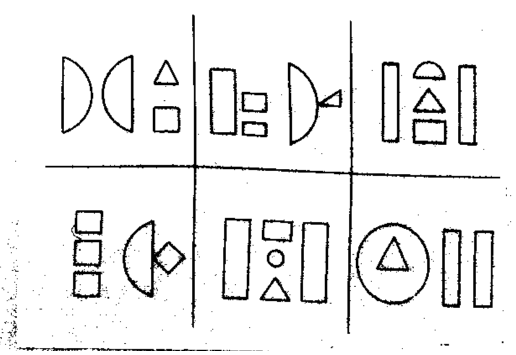
Design A



Design B



Design C



Design D

Appendix C

Participants Information Sheet

Victoria University

Department of Psychology

Invitation to Participate in a Research Study:

Development of a scoring system for an alternative form of the Visual Reproduction subtest of the Wechsler Memory Scale – Revised.

My name is Daniela Petrov and I am undertaking the Doctor of Psychology (Clinical Neuropsychology) degree at Victoria University. I am being supervised by Dr Peter Dowling in the Department of Psychology at Victoria University, St. Albans.

I am currently involved in a study that looks at the development of a scoring system for an alternative memory test that can be used in the assessment of memory functioning.

Memory tests are important tools for understanding the effects of brain injury or disease on memory functioning. Neuropsychologists often use these memory tests to determine if there are any improvements in client's memory functioning due to rehabilitation or a treatment program.

I would like to invite you to be a part of this study. Participation in this study will involve you doing a number of tasks that will mostly look at memory, learning and intellectual functioning. A brief questionnaire about your age, education, medical history and your general well-being will also be included. The tasks will be conducted over two sessions scheduled one week apart, and each session will last approximately an hour. These sessions will be conducted in the comfort of your own home at at a convenient time for you.

Like any neuropsychological assessment you may experience some fatigue during testing. If you are feeling tired, breaks will be available for you to take as you request. Your participation in this study is voluntary and you are free to withdraw at any time, without any obligations.

All information that I obtain will be kept strictly confidential and only group results will be published without any names or other identifying information.

Should you have any questions or queries regarding this study please do not hesitate to contact Dr. Peter Dowling or myself through the Victoria University Psychology clinic on 9#######.

Thanking you for your time,

Daniela Petrov

Appendix D

Consent Form

CONSENT FORM

CERTIFICATION BY PARTICIPANT

I, ____
of ____

certify that I am at least 18 years old and that I am voluntarily giving my consent to participate in the study entitled:

> Development of a scoring system for an alternate form of the Visual Reproduction subtest of The Wechsler Memory Scale- Revised.

being conducted at Victoria University of Technology by:

Daniela Petrov
Dr. Peter Dowling

I certify that the aims of the study, together with any risks and procedures to be carried out have been fully explained to me. I freely consent to participation involving the use of these procedures outlined in the ‘invitation to participate’ statement.

I certify that I have had the opportunity to have any questions answered and that I understand that I can withdraw from this study at any time and that this withdrawal will not jeopardise me in any way.

I have been informed that the information I provide will be kept confidential at all times.

Signed: .... }

Witness other than the researcher (as appropriate) } Date: ....

....}

Any queries about your participation in this project may be directed to the primary researcher (Name: Dr. Peter Dowling ph. 9#######). If you have any queries or complaints about the way you have been treated, you may contact the Secretary, University Human Research Ethics Committee, Victoria University of Technology, PO Box 14428 MCMC, Melbourne, 8001 (telephone no: 03-9#######).