5-2010

Cluster analysis of the TOMAL standardization sample

Nicholas Shizuo Thaler
University of Nevada Las Vegas

CLUSTER ANALYSIS OF THE TOMAL STANDARDIZATION SAMPLE

by

Nicholas Shizuo Thaler

Bachelor of Science
University of California, San Diego
2004

Master of Arts, Clinical Psychology
California State University, Northridge
2006

A thesis submitted in partial fulfillment
of the requirements for the

Master of Arts Degree in Psychology
Department of Psychology
College of Liberal Arts

Graduate College
University of Nevada, Las Vegas
May 2010

UNLV
UNIVERSITY OF NEVADA LAS VEGAS

THE GRADUATE COLLEGE

We recommend the thesis prepared under our supervision by

Nicholas Shizuo Thaler

entitled

Cluster Analysis of the TOMAL Standardization Sample

be accepted in partial fulfillment of the requirements for the degree of

Master of Arts in Psychology

Daniel N. Allen, Committee Chair

Brad Donohue, Committee Member

Murray G. Millar, Committee Member

Chad Cross, Graduate Faculty Representative

Ronald Smith, Ph. D., Vice President for Research and Graduate Studies
and Dean of the Graduate College

May 2010

ABSTRACT

Cluster Analysis of the TOMAL Standardization Sample

by

Nicholas Shizuo Thaler

Dr. Daniel N. Allen, Examination Committee Chair
Professor of Psychology
University of Nevada, Las Vegas

Children experience natural cognitive changes as they grow older, with more rapid memory development when they are younger. The purpose of the present study was to investigate the expected normal variation in youth using the Test of Memory and Learning (TOMAL; Reynolds & Bigler, 1994). The TOMAL is a broad-band verbal and nonverbal memory battery that has been standardized on youth aged 5-19, providing a good source of information on memory development in healthy children. Cluster analysis analyzed 1121 children selected from the TOMAL standardization sample and identified homogeneous profile subtypes of memory and learning. In addition, three age ranges were determined via statistical analysis and subsequent cluster analysis were run on each of these ranges. Results found that a 5-cluster solution was optimal for the entire sample, a 4-cluster solution for the 5-8 year age group, a 5-cluster solution for the 9-11 group, and a 6-cluster solution for the 12-19 group. The 5-8 year age group exhibited variation in level of performance on TOMAL index and factor scores while other groups exhibited variation in both level and pattern of performance on these scores. These findings indicate a clear pattern of expected memory variation in normal children, with differential performance in verbal memory, nonverbal memory, attention/concentration, working memory, and spatial memory

domains. Results also indicate that memory performance grows increasingly complex and diversified in older children as compared to younger children.

ACKNOWLEDGMENTS

I'd like to thank Dr. Daniel Allen for being a wonderful mentor and friend and Dr. Marta Meana for being the best clinical supervisor a graduate student could ask for. I'd also like to thank my lab mates Erik, Sally, Griffin, Cris, and Carol for all of their hard work and friendship. My fellow cohort students including Stacy, Rachel, Kerri, Kelsey, Harpreet, and Johanah have also been wonderful companions whom I share this journey with. I also thank Ani, Dave, Adam, Justin, and Yashar for keeping me grounded and providing me with a link outside of academia. Finally, I'd like to acknowledge my parents, Agnes and Richard, my brother Ben, and my other family including Mitzi, Milt, Koka, Keiko, and Dwight for providing me with all the opportunities that I have today.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

A well established memory battery is an important tool for neuropsychologists, as memory deficiency is a frequent symptom exhibited across a wide array of neurological disorders (Reynolds & Bigler, 1997). Traumatic brain injury (TBI), attention-deficit hyperactive disorder (ADHD), schizophrenia, bipolar disorder, dementia, and learning disabilities are a few of the clinical disorders that exhibit significant memory disturbances. Memory has been called the "cornerstone of cognition" (Reynolds & Bigler, 1996), and encapsulates how we as individuals process and interact with the world (Siegel, 2001) by encoding our interactions into our brain in ways that affect our emotions and behaviors. The complex role that memory plays into our sense of identity and thought makes it particularly vulnerable to neurological disorders. Therefore, a careful and thorough assessment of broad and specific memory functioning can be instrumental to informing clinicians about the overall condition of a patient.

Four predominant neurocognitive memory batteries are currently available for children and adolescents. These memory batteries are the Wide Range Assessment of Memory and Learning (WRAML; Sheslow & Adams, 1990) and its second edition the Wide Range Assessment of Memory and Learning-2 (WRAML-2; Sheslow & Adams, 2003), the California Verbal Learning Test for Children (CVLT-C; Delis, Kramer, Kaplan, & Ober, 1994), the Children's Memory Scale (CMS; Wechsler, 1995), and the Test of Memory and Learning (TOMAL; Reynolds & Bigler, 1994), now in its second edition as the Test of Memory and Learning-2 (TOMAL-2; Reynolds & Voress, 2007). The TOMAL in particular has received much recognition as a comprehensive measure with subtests that encompass both broad-based and narrow-based aspects of memory (Ferris & Kamphaus, 1995; Lowe, Mayfield, & Reynolds, 2003). The measure is composed of fourteen subtests that cover a wide range of memory abilities that tap into verbal and nonverbal memory, attention, and learning.

Because of TOMAL's comprehensive assessment of memory, it a widely used measure in assessment. It has been used to detect brain dysfunction in ADHD (Morrison, 2006), TBI (Lowther & Mayfield, 2004), and reading disabilities (Howes, Bigler, Lawson, & Burlingame, 1999). The TOMAL has been shown to have high reliability and norms are available based on a large, representative standardization sample (Reynolds & Bigler, 1994b). The success of this battery has led the authors to revise it for use with adult populations. The first standardization sample was restricted to an age range of 5-19, and until recently this limited the TOMAL only for child and adolescent populations. The TOMAL-2 (Reynolds & Voress, 2007) has extended the standardization sample to include adults, increasing the age range to 5-59 years. Early psychometric testing of the TOMAL-2 is promising, with reliability coefficients and factor structures similar to that of the TOMAL (Reynolds and Voress, 2007).

Psychological measures that have strong reliability and validity can be further analyzed using cluster analysis, a descriptive technique that categorizes a heterogeneous pool of participants into smaller, homogenous groups or clusters (Romesburg, 1984) that are partitioned by specific attributes that are of interest to the researcher. These clusters can be compared to other demographic or performance variables to confirm their convergent and external validity. This technique has been used on the standardization samples of several measures including the Wechsler Adult Intelligence Scale-Third Edition (WAIS-III; Wechsler, 1997), the Wechsler Intelligence Scale for Children-Third Edition (WISC-III; Wechsler, 1991), and the adult and child versions of the CVLT (Donders, 2008; Donders, 1999). The clusters formed from these measures exhibit significance differences in both level and pattern of performance (Donders, 1996; Donders, Zhu, & Tulsky, 2001; Lange, 2007), suggesting that these clusters represent subtypes of ability in the constructs measured by the batteries.

Cluster analyses have been performed on both normal and clinical populations. Clusters identified in normal populations provide valuable information regarding variability in patterns of performance that occur in the general populations which may be useful in determining whether a particular profile represents individual differences, or is indicative of abnormal cognitive processes. Researchers have also ran cluster analyses on samples exhibiting specific neurological disorders. For example, TBI patients typically exhibit a wide variability of scores on measures of intelligence and memory. Grouping these scores into identifiable clusters may parse the heterogeneity of symptoms expressed in TBI into smaller, more homogeneous subgroups. Identifying homogeneous groups within a heterogeneous clinical population can aid clinicians in diagnosing the severity of a disorder and designing individual treatment plans for patients who fall within certain clusters. To date, researchers have successfully identified clusters in TBI patients that differed across levels and patterns of performance on measures such as the WISC-III and the CVLT (Heijden & Donders, 2003; Donders, 2006). In turn, these clusters significantly differed across demographic, clinical, and behavioral variables, confirming the grouping of symptoms within the subgroups.

With regard to the TOMAL, a cluster analysis was conducted on children with TBI (Allen, Leany, Thaler, Cross, & Mayfield, in preparation) and five clusters were identified as the optimal solution, three that differed on levels of performance and two that had different patterns of performance in the verbal or nonverbal domains. However, no cluster analysis study yet exists for the standardization sample of the TOMAL. The purpose of the present study is to address this gap in the literature and examine the best fit of clusters that emerge on the TOMAL standardization sample. Furthermore, because the TOMAL assesses young children and adolescents and changes in developing cognitive abilities occur as children develop, three separate age groups were examined (i.e. 5-8, 9-11, 12-19) in order to evaluate the similarities and differences of the clusters between different age ranges.

CHAPTER 2

LITERATURE REVIEW

Before discussing the current research on the TOMAL, it is important to establish why memory is a domain of interest within the field of neuropsychology. Specifically, we will start with the concept of memory, and discuss its natural development and progression with aging. Following this, we will explore the expected genetic variation of memory performance in clinical and non-clinical populations and memory's potential role as an endophenotype. Cluster analysis will then be discussed, followed by a review of articles that used cluster analysis to research the psychometrics of neuropsychology measures with particular emphasis on studies of children. Next, the TOMAL and some studies on its validity are discussed. Finally, the review will conclude with studies that have used cluster analysis specifically for the TOMAL.

Memory: Models and Progression

Long ago, psychologists identified memory as a key aspect of cognition. Some have suggested that memory is the cornerstone of our identity, granting us the powers to reason and plan, to reflect and act (Reynolds & Bigler, 1994). It is small wonder that memory is a common focus of rehabilitation and training for those who have sustained brain damage (Prigatano, 1990). Research has confirmed that memory is a multifaceted construct, with many subtypes that function independently from each other (Siegel, 2001). The theories behind memory functioning are not fully cohesive, but the aggregate of research suggests the presence of two constructs of memory that can be split into verbal and nonverbal domains (Reynolds & Bigler, 1994). It is generally agreed that verbal and nonverbal performance are lateralized in the two hemispheres of the brain, with the right hemisphere typically responsible for processing of nonverbal information, and the left hemisphere for processing of verbal information.

How does our brain encode, store, and retrieve our memories? Atkinson and Shiffrin (1969) have proposed that our brain is composed of three stages or stores, including the sensory store (SS), short-term memory (STM), and long-term memory (LTM). Information in the SS is an exact representation of the environment and only those aspects that are most relevant are selected by attention for additional processing (Thompson, 2000). Relevant information passes from the SS to the STM store. The STM store is a limited capacity store that can hold 7+-2 pieces of information (Miller, 1956). Information is maintained in the STM store through rehearsal, but quickly fades once it is no longer attended to, at an estimated 18 seconds (Peterson & Peterson, 1959). Information that is sufficiently rehearsed is ultimately encoded into the LTM store. The LTM hypothetically possesses a limitless capacity and information is maintained indefinitely in this store (Bahrick, 2000).

The Atkinson and Shiffrin model has proven very useful in understanding basic components of the memory processing system, but in its initial form probably was too simplistic to explain the complexity of memory systems, and as a result has been subsequently elaborated by other investigators. For example, Baddeley and Hitch (1974) detailed on the STM store with the concept of working memory (WM). WM serves as a type of holding cell that temporarily makes a memory easily accessible for use. WM also may be identified as the first line of conscious awareness. Baddeley's model of WM describes three components: a central executive component and two "slave" systems comprised of a phonological loop system and a visuospatial sketchpad system. Later, Baddeley (2000) added a fourth component, the episodic buffer system which links information across the other systems and to LTM storage. Milner and colleagues (1998) posited that WM does not rely on genetic transcription and translation, but rather with immediate chemical alternations within the synapses. Therefore, WM temporarily stores memory which fades rapidly. Though useful for transient recordings, such as remembering a phone number long enough to use it, working memory must be encoded into long-term memory if it is to be held for storage.

Just as WM has been distinguished as a process separate from general STM storage, LTM has also been subdivided in a number of ways, such as verbal vs. nonverbal and declarative vs. procedural. Declarative memory stores facts and has been further divided into semantic and episodic memory. Declarative memory refers to memories that we are consciously aware of and can explicitly recall (Tulving, 1984). Semantic declarative memory refers to knowledge of specific facts free from any personal time or place (Thompson, 2000). An example of a semantic memory is knowing that a dog is a type of animal. Episodic memory refers to factual knowledge of an event that is personally relevant to the individual in terms of time and place. Remembering the dinner you had last night would be an episodic memory. Declarative memory is primarily stored in hippocampus, located at the medial temporal lobe of the brain.

Nondeclarative memory, also called implicit memory, is more subconscious in nature (Thompson, 2000). Procedural memory, a form of nondeclarative memory, refers to our ability to learn specific skills and abilities that may not be consciously accessible. For example, learning the chords to a guitar or how to ride a bike involves procedural memory. In contrast to declarative memory which is stored primarily in the hippocampus, procedural memory is stored in the amygdale, cerebellum, striatum, and motor cortex as well as in the hippocampus. Habituation and sensitization, two forms of procedural learning, are nondeclarative and present in both animals and humans. Habituation as proposed by Solokov (1963) refers to our decrease in response to repeated stimulation. When we perceive something new or alien we are more likely to attend to it then when we perceive something we've seen many times before. Sensitization refers to an increase in response to repeated stimulation, usually stimulation that is particularly strong. Both habituation and sensitization are present in primitive nervous systems and may be viewed as some of the earliest mechanisms of procedural memory.

Associative memory is another form of implicit memory that involves associative learning, which is the forming of associations among stimuli and our responses to these stimuli (Thompson, 2000). Much of the science behind behaviorism involves associative learning and memory, including classical and operative conditioning. Studies by researchers (i.e. Davis et al., 1987) have identified the amygdala as a key structure involved in associative learning and memory.

LTM storage undergoes a process called cortical consolidation (Abel et al., 1998). Though not well understood, it is believed that rapid eye movement sleep may contribute to the consolidation process (Winson, 1993). Once a memory has been consolidated, it may be considered much more permanent for the individual. The permanent nature of LTM at this stage suggests that permanent changes have occurred on the cellular level (Thompson, 2000). Volkmar and Greenbrough (1972) have found that a wide variety of learning experiences increases dendritic branding in neurons in rats, some of the first evidence of neuronal structural changes associated with learning experiences. Long-term potentiation (LTP) has been hypothesized as one of the chief mechanisms underlying the cellular changes associated with memory. The mechanism behind LTP is that synapses may have an increase in excitability due to activation of NMDA CA2+ channels (Thompson, 2000). When glutamate from a presynaptic cell binds to AMPA receptors in the postsynaptic cell, a chemical reaction takes place that allows calcium to flow into the cell. Researchers such as Lynch and Baudry (1984) have provided evidence that the cellular processes triggered by calcium entry can actually alter the shapes of synapse proteins permanently and make the synapses even more receptive to glutamate. These structural changes associated with frequent stimulation of neurons may be one of the mechanisms behind LTM storage.

The process of taking in sensory stimuli, holding and manipulating the sensory memory into information in the STM or WM store, and encoding the information into LTM store involves several systems, and damage to any one of them can cause memory disruption. For example, damage to the hippocampus inhibits the brain’s ability to encode new material into long term memory, resulting in anterograde amnesia. Previously stored memories remain intact, allowing the individual to recall past memories, but new information is not encoded into LTM store (Miller, Bigler &

Adams, 2001). In another example, damage to the left hemisphere has been associated with impairment in verbal memory, while damage to the right hemisphere with visual-spatial memory.

Memory Development

Infants are born with their implicit memory functioning intact. Explicit memory does not develop until the brain further matures, usually by the second year of life. This accounts for the phenomenon known as "childhood amnesia," or the inability to recollect anything before the age of three. As explicit memories develop, children achieve a self-reference in their everyday actions and environment. This creates a sense of continuity and mental time travel, or remembering events in the past or predicting events in the future (Siegel, 2001). Ultimately, autobiographical memory is dependent on the development of the hippocampus and the orbitofrontal regions, which may not fully mature until the age of five.

Once children are of five years of age, they are generally regarded as having the same capacity to form new memories and recollect past ones as an adult. Most memory batteries designed for children and adolescents start at this age, including the TOMAL. Much of the research on memory development focuses the formation of memory strategies in children. For example, as children go to school, their strategies for memorizing semantic facts shift from a more passive stance to an active one. Research suggests that the environment has an influence on memory strategies. For example, Coffman and colleagues (2008) conducted a longitudinal study on 1st grade children. They found that teachers who emphasized mnemonic strategies as part of their teaching style had children who had better memory skills compared to other teachers. Other similar studies confirm that most children acquire memory strategies in elementary school, and that teachers and other environmental influences affect the development of these strategies.

As children grow older they develop enhanced STM and WM skills, with particular gains in perceptual analysis, construction and maintenance of memory traces, retention of order information, and complex working memory functioning. A comprehensive review by Gathercole (1998) details the natural developmental processes of several memory components. The phonological loop portion of working memory appears to be intact in young children, but children do not subvocally rehearse units of information until around 7 years of age (Gathercole & Hitch, 1993) after which this type of memory increases rapidly and then steadies around age 9. Visuospatial memory appears to develop rapidly up until around age 11, at which point it steadies through adulthood (Gathercole, 1998). On the other hand, complex working memory tasks that load more heavily on the central executive appear to develop steadily through adolescence and adulthood.

Supporting this, another study by Gathercole (1999) found that digit span, non-word repetition, and visuospatial memory all dramatically increased from ages 2-8 and then gradually increased from ages 9-16 (see Figure 1). The exception to this was a complex form of working memory that showed steep improvement through age 16. WM is held in the frontal lobe, a region of the brain that continues to develop through adolescence. In contrast, temporal, parietal, and occipital lobes reach adult levels of activity by three to six months of age. In another study looking at the relationship between WM and age, Swanson (1999) conducted a study on the development of verbal and visuospatial WM across several age groups. He found a linear relationship between WM and age and that increased performance related to memory access and storage demands. He argued that increased controlled attentional resources as children developed contributed to improvement of WM functioning.

In summary, as the brain matures, so does the ability to form and retain explicit memories. This process is complete around the age of five. From here, children develop memory strategies that are influenced by both the environment and their own natural abilities, enabling them to recollect more complicated memories, including the detailed semantic facts often required in educational settings. Phonological STM stores rapidly increase from ages 2 through 8 while visual-spatial STM increase up to age 11, and both steadily improve through adolescence. On the other hand, complex WM tasks rapidly increase throughout adolescence and adulthood. Memory peaks in young adulthood, and slowly begins a decline in middle age. As our study focuses on children and adolescents, it would appear from Gathercole's (1998; 1999) findings that age 9 and age 12 may be appropriate cutoffs when distinguishing the memory processes of younger children from older children.

Memory and Genetics

Memory is expected to vary in normal populations because like many cognitive networks, memory is a heritable trait (Gazzaniga, 1995). It is well established that memory variation is related to other cognitive processes such as language development (Gathercole & Baddeley, 1989) and speech production (Spiedel, 1991). Research has also suggested that genetic predisposition is a contributing factor to memory variation in normal, healthy individuals (Ando, Ono, &

Wright, 2001; Kremen et al., 2007) One protein that has gained considerable attention in the mental health field is Catechol-O-methyl transference (COMT), a gene located on chromosome 22 that catabolizes dopamine (Payton, 2006). COMT activity is critical in the prefrontal cortex, a region of the brain with many dopaminergic pathways. The prefrontal cortex is a region responsible for complex executive processing which in turn is responsible for working memory tasks (Baddeley, 2000). It is likely that variability in COMT activity plays a role in strategic memory processing among normal individuals (Raz, Rodrigue, Kennedy, & Land, 2009).

Another gene that may play a role in memory is the brain-derived neurotrophic factor (BDNF) gene located in chromosome 11. BDNF is a neurotrophin located in neurons and glia that is responsible for the proliferation, differentiation, and survival of neurons (Binder & Scharfman, 2004; Murer, Yan, & Raisman-Vozari, 2001). Studies in rats and other animals have found that BDNF plays a role in learning and memory (Finkbeiner, 2000; Lynch, Rex, & Gall, 2006). Other studies have shown that adults with reduced BDNF expression have poorer verbal memory than adults with higher rates of BDNF expression (Dempster et al., 2005; Ho et al., 2006).

Finally, Apolipoprotein E is a gene found on chromosome 19 that controls lipid transport and high levels of this gene is associated with Alzheimer’s dementia (Corder et al., 1993) and declines in episodic memory in healthy adults (Tupler et al., 2007; Packward et al., 2007). However, these results are not always consistent as another study linked Apolipoprotein E with increased episodic memory in young adults (Mondadori et al., 2007). Therefore, this gene may have an interaction with age in its role in memory performance.

The available literature on the relationship between memory and genetics suggests that researching memory's role as an endophenotype is the next step in identifying the etiology and course of many pervasive psychological disorders. Brzustowicz and Bassett (2008) have espoused the need for relevant trait markers, or endophenotypes that can distinguish neurological disorders from other pathologies and conditions. Endophenotypes are classified as a unique form of biomarker that clusters behavioral symptoms into stable phenotypes that have a genetic source (Gottesman & Gould, 2003). Examples of endophenotypes include sensory gating, attention, mental flexibility, and working memory. For a biomarker to be included as an endophenotype, it must meet five criteria: 1) it is linked to an illness within the population, 2) it is heritable, 3) it is stable within an individual, whether or not symptoms to its corresponding illness are present, 4) families of individuals with the illness carry the endophenotype at a rate higher than the general population, and 5) the endophenotype and its associated illness co-segregate.

Memory has been proven to be a reliable endophenotype for a number of neurological disorders. For example, Frantom, Allen, and Cross (2006) gave a comprehensive battery of tests to patients with Bipolar Disorder I, their first degree relatives, and normal controls. The measures assessed a multitude of cognitive domains including as executive functioning, attentional control/processing speed, working memory, verbal learning and memory, and visual learning and memory. Frantom and colleagues found that the Bipolar I group and their first degree relatives had marked deficits on measures that tasked executing functioning, visual learning ability, and motor domains, with the relatives performing at an intermediate level compared to the bipolar group and the controls. These findings suggest that the domains that differed across the three groups are indicative of endophenotypes unique to bipolar individuals. If clinicians can reliably distinguish bipolar disorder with the use of such neurological tests which identify discrepancies in specific endophenotypes, we would have an alternative approach to assessment that may provide further information on the nature and severity of a disorder.

Another study conducted by Tabares-Seisdedos and colleagues (2007) investigated cognitive endophenotypes that may predict long-term outcomes for schizophrenic and bipolar patients in a yearlong longitudinal study. Subjects were given tests that measured eight neurocognitive domains, specifically executive functioning, verbal working memory, verbal memory, visual memory, visual-motor processing speed, vigilance, motor speed, and vocabulary. Researchers found that deficiencies in verbal memory and motor speed predicted schizophrenic patients, while deficiencies in visual/motor processing predicted bipolar disorder patients. Furthermore, the one year follow-up assessment showed that impairments in executive functioning were the best predictor of a poor prognosis for the bipolar patients.

Other research suggests that impulse-aggressive markers may be a link for suicidal patients (Zouk et al., 2007), deficiencies in nonverbal memory are associated with obsessive-compulsive disorder (Rao et al., 2008), and deficiencies in state regulation and delay aversion have been linked to attention-deficit hyperactive disorder (ADHD; Doyle et al., 2005). Though the research is still preliminary and inconclusive, memory has come up again and again as key endophenotypes that may be indicative of psychological disorders. Furthermore, different *types* of memory have predicted different disorders; for example, bipolar disorder is generally associated with deficits in visual memory, and schizophrenia with deficits in verbal memory.

The above literature identifies memory as a genetically-linked endophenotype that has expected variable patterns in clinical populations. In addition to clinical populations, genetically driven variation of memory abilities in nonclinical populations is expected. This variation leads to heterogeneous performance of memory batteries in representative standardization samples. Cluster analysis is a taxometric method that classifies objects into homogenous and discrete subgroups (Romesburg, 1984). Therefore, it is an ideal technique for parsing the expected variation in the TOMAL standardization sample into smaller subgroups. We now turn to an overview of cluster analysis.

Cluster Analysis

Cluster analysis is used in many disciplines, including biology, geology, anthropology, and marketing (Tryon, 1939). Before cluster analysis can be performed, a set of objects must be arranged in a data matrix. In most cases, the columns of the matrix represent the individual objects, while the rows represent a set of determined attributes that each object may or may not possess. For example, an archaeologist may be interested in determining the evolutionary link of an unspecified set of bones. The archaeologist can identify several physical, chemical, and other attributes of these bones and arrange them as rows on a matrix. Then, the bones and other bones that have already been classified are laid out as columns. Cluster analysis uses a variety of mathematical methods to determine which classified bones are the most similar to the unknown bones, based on the determined attributes (Kaufman & Rousseauw, 1990).

Romesburg (1984) outlined three research goals that cluster analysis can answer. The first goal is to create a question to be tested later. Creating a question is relatively simple, as the researcher can simply run a cluster analysis on a data matrix and observe what clusters form together. Though it would be irresponsible to draw any conclusions without a hypothesis, it is appropriate to further investigate any interesting patterns that emerge in subsequent studies. The second goal is to create a hypothesis. The researcher already has a question framed when running the analysis, but no testable hypothesis. Any patterns that emerge may answer the question and open up the possibility of a hypothesis. Finally, cluster analysis can be used to test a hypothesis. Typically, previous studies that may or may not have already used cluster analysis have presented evidence of a clear, testable hypothesis. The hypothesis must be made *a priori* and any conclusions must be directly related to the hypothesis. Most of the literature on psychometric measures already has a firmly developed hypothesis. In the current study, cluster analysis will likewise be used to address this question.

This technique can be a useful tool in psychometrics, as the attributes are already predetermined in the form of set items on a test, while the objects of interest are generally the participants. Cluster analysis can be used to see how different people group together into clusters based on how they answered the items. These newly formed clusters can then be treated as different groups, and compared on variables that were not included in the cluster analysis in order to determine if the groups differ in meaningful ways that would suggest a valid taxomony. For example, researchers interested in TBI have administered the WISC-III to a TBI sample, and then run a cluster analysis to determine if there are subgroups or clusters of patients that are differentiated by their performances on the WISC-III scales (Donders & Warschausky, 1997). Each of these clusters represents TBI patients who have scored similarly across the WISC-III. The clusters themselves have little meaning, until they are compared to variables external to the cluster solution. For example, one WISC-III cluster may be linked with longer periods of unconsciousness and poorer outcomes, while another cluster may have relatively mild injuries with better outcomes. Indeed, there is a proliferation of literature on investigating the validity of neuropsychological measures' clusters that emerge in clinical populations (Goldstein, Allen, & Seaton, 1998; Mottram & Donders, 2006; Seaton, et al., 1999).

Resemblance Coefficients

Once a researcher has put together a data matrix, the researcher determines how to analyze the data by choosing a resemblance coefficient. There are several resemblance coefficients to choose from, but each coefficient is either a similarity or dissimilarity coefficient. This dichotomy simply expresses the direction of the data; when using a similarity coefficient, larger values indicate higher similarity between two objects while the opposite is true with a dissimilarity coefficient. A review of the literature in psychology indicates that the Euclidean distance coefficient is the most common distance measure in published studies (Clatworthy et al., 2005) which finds the least distance between two objects via Euclidean geometry. This coefficient can easily be visualized when only two attributes are compared across the objects. These two attributes are treated as coordinates on a two-dimensional plane, and a point on the plane represents an object. The Euclidean distance coefficient calculates the linear distance between objects by using the Pythagorean theorem. Therefore, the farther two points are, the more dissimilar the represented objects are from each other.

In most matrices, objects are compared across more than two attributes. A three-attribute cluster analysis can be envisioned as a three-dimensional space, but higher attribute analyses cannot be pictured as easily. Nevertheless, the principle remains the same: the Euclidean distance coefficient calculates the overall distance that two objects are from each other in a hypothetical space. These distances are placed on a new matrix called the resemblance matrix, with which researchers can determine the similarity between individual objects. However, how objects actually combine to form clusters is determined by a second technique called the clustering method.

Like with distance coefficients, the researcher determines the optimal clustering method and there are many methods that he can select (Kaufman & Rousseauw, 1990). Clustering methods can be hierarchical or partitional in nature. Hierarchical methods are the preferred form for most researchers, as they build dendograms, or trees, which are visual representations of the clusters. The majority of the literature in neuropsychology uses Ward's minimum variance clustering method (Ward, 1963), which is also the second most used clustering method across all scientific fields (Romesburg, 1984). Like all hierarchical methods, Ward's method is agglomerative, building clusters from individual objects and combining clusters based on their similarity to each other until the final cluster, which encompasses all the data, is formed. This final cluster can be visualized as the "trunk" of the tree, which in turn breaks into smaller and smaller branches, while the tips of the tree represent the original objects. Ward's method calculates similarity by using a sum-of-squares calculation to see which two items exhibit the least variance when combined into a hypothetical "average." All cluster combinations are compared at each level of the tree, and a new cluster is formed each time the smallest variance is found. This continues until all objects are formed into one unifying cluster.

Another hierarchical clustering method worth noting is the two-step clustering method, which has the advantage of automatically selecting the number of clusters and handling categorical as well as continuous variables (Bacher, Wenzig, & Vogler, 2004). The two-step method clusters individual cases into small sub-clusters, and then clusters these sub-clusters into the cluster solution. In large datasets with only continuous variables, such as the dataset in this study, the Euclidean distance coefficient is used.

Once the dendogram is fully formed, researchers must determine where to "cut" the tree, or where the optimal cluster solution is found. The optimal cut is subjective, but typically a smaller cluster solution is preferred over a larger one. Romesburg (1984) recommends that the tree should be cut where the clusters are maximally related to other variables of interest. Therefore, cutting the tree in different ways may produce different results, and the one that fits the proposed hypothesis the best should be selected.

There may be some unforeseen complications that emerge from the data. Chaining is a term used to describe a cluster that repeatedly merges with individual objects, much like a black hole absorbs random pieces of debris (Anderson, 1973). Ideally we would want objects to clump into several smaller clusters and only merge together into the single cluster at the very end of the analysis. With chaining, it is more difficult to determine the similarity of objects as each object is added one at a time to a single, growing cluster. Another complication can emerge when the dendogram does not accurately represent the data matrix (Romesburg, 1984). This can occur because clustering methods mathematically calculate the similarity of objects using formulas that do not exactly match the actual similarity in Euclidean space (or, if another coefficient is used, whatever is determined to represent similarity among objects). Researchers typically avoid this problem by calculating the cophenetic correlation coefficient, a Pearson's correlation between the actual data matrix and the proposed matrix formed from the dendogram. Correlations that are greater than .80 indicate that the distortion between the matrix and the dendrogram is not severe.

Interpreting the Clusters

Cluster analysis is exploratory in nature. While the data matrix includes attributes that are selected based on their ability to provide a meaningful classification of the objects, there are likely several other attributes and associations among the clusters that are unknown (Aldenderfer & Blashfield, 1984). Therefore, after deriving the clusters, the researcher must subsequently establish the clusters' stability and validity.

Depending on the resemblance coefficient, the clustering method, and the sampled data, clusters may vary on separate analyses. It is important to establish the stability of formed clusters. As already mentioned, clustering methods can be hierarchical or partitional in nature. Hierarchical clustering methods are preferred in the available literature, but partitional methods are often used to confirm the stability of these clusters (McIntyre & Blashfield, 1980). Partitional clustering methods determine all clusters simultaneously, as opposed to the “bottom-up” or “top-down” trees found in hierarchical models. The K-means iterative classification process is a partitional method often used in tandem with Ward’s method. The K-means method either randomly or purposefully defines starting points, or “centroids” for a predetermined number of clusters (MacQueen, 1967). Individual objects are assigned to the cluster center with the most similarity. The centers of the clusters defined by Ward’s method are assigned as the starting points for the K-means method. The extent to which the K-means method and Ward’s method agree establishes the clusters’ stability.

Establishing the validity of a cluster solution depends on the nature of the research. If the researcher’s goal is to separate a large data set into homogenous groups classified by size or color, then the validity of the clusters is intrinsic to their physical properties. In the case of psychometric data, the formed clusters may only be the starting point of further data analyses. Participants may respond to psychological measures in ways that can be classified into smaller and more homogenous subgroups, but this in itself is not particularly informative to the researcher. Rather, the researcher must compare these clusters on a wide range of other variables and see if they significantly differ. With careful and repeated study, clusters can eventually be confirmed as having external validity with real world variables, enhancing their utility for research.

Cluster Analysis and Neuropsychology Measures

Clinical Samples

As previously mentioned, cluster analysis of neuropsychological variables has been accomplished with both clinical and nonclinical populations. The goal in examining clinical populations is to identify pattern and level of performance across neuropsychological measures that might provide insight into the differential involvement of both structures and circuits that contribute to disease expression. Studies of psychiatric disorders such as schizophrenia have found that cluster solutions based on cognitive tasks (e.g. Wechsler intelligence scales) can predict demographic variables. For example, Goldstein, Allen, and Seaton (1998) derived a four-cluster solution based first on the subtests of Wechsler Intelligence scales and then on a battery of neuropsychological tests (i.e. Halstead Category, the Trail Making Test, etc.) taken by 221 schizophrenia patients. Cluster membership was found to have better external validity with the Wechsler scale solution. Another study by Seaton et al. (1999) replicated the four-cluster solution found in a different sample of schizophrenia patients and compared the clusters across demographic and clinical variables. The authors found that cluster membership had significant differences across years of education, occupational functioning, and years of illness, suggesting that the clusters were sensitive to the variability of adaptive functioning in schizophrenic patients.

Studies of traumatic brain injury (TBI) have also shown variability in pattern and level of performance. Donders and Warschausky (1997) ran a cluster analysis on a sample of children with TBI using WISC-III index scores. The researchers found that a four-cluster solution was the optimal fit, with three clusters differing across levels and a fourth cluster by patterns of performance. The three levels of performance ranged from “high-average” to “below-average.” The fourth cluster had low scores on the Processing Speed Index (PSI) and the Perceptual Organization Index (POI) scores compared to the other index scores. Donders and Warschausky subsequently investigated the validity of the clusters by comparing them across other variables. There were no statistically significant differences due to gender, type of injury, age of testing, or ethnicity. However, the high-average cluster did come from a higher socioeconomic background, while the low POI and PSI cluster came from a lower one. Severity of injury as determined by Glasgow Coma scores (GCS) and length of time in coma also distinguished the clusters. Finally, right-cerebral lesions were more associated with the low POI and PSI cluster compared to other clusters. The authors concluded that this cluster in particular was meaningful, as it directly was linked to overall poor WISC-III performance and markedly discrepant POI and PSI index scores. They also note that no “high functioning” PSI cluster was found, compared to the standardization sample. Such findings suggest that TBI is particularly damaging to processing speed, a fact that is consistent with other studies on pediatric TBI (e.g. Bawden, Knights, & Winogron, 1985).

Heijden and Donders (2003) have examined clusters in WAIS-III TBI profiles in adults. The WAIS-III has reliable factor index scores that are useful for interpretation (Kaufman, Lichtenberger, & McLean, 2001). Heijden and Donders ran a cluster analysis on 166 adult patients with TBI. The sample presented a three-cluster solution as the optimal fit, with individual clusters varying on levels of performance. Levels of education and severity of injury were modulating variables that differentiated the clusters. Though the authors did not find a hypothesized fourth cluster that had a particularly weak PSI score, they did note that all three cluster profiles had relatively low PSI scores compared to other index scores. The authors also addressed the lack of a low POI score, which had emerged on a fourth cluster on WISC-III TBI profiles (Donders & Warschausky, 1997). They posited that WISC-III PO subtests have greater a reliance on processing speed compared to the WAIS-III PO subtests, accounting for the lack of a low POI score in any of the WAIS-III clusters. In essence, processing speed still remains particularly sensitive to TBI in adulthood, though perhaps not as much as in childhood.

Mottram and Donders (2006) used cluster analysis on a pediatric sample of TBI patients. The researchers posited that any differences that emerge in this cluster solution compared to that of the CVLT-C standardization sample would suggest unique profiles in TBI. Unlike with the standardization sample, Mottram and Donders (2005) selected four empirically derived factors of the CVLT-C as the attributes rather than using subtests. Three of these factors match the variables that Donders (1999) used (Attention Span, Learning Efficiency, and Inaccurate Recall) while a fourth (Delayed Recall) combined the other two variables. The sample consisted of 175 boys and girls with pediatric TBI. A four-cluster solution emerged in which three clusters differed on levels of performance (high average, average, low average) and a fourth on patterns of performance. The fourth cluster was distinguished by having low scores on all factors with the exception of a normal Inaccurate Recall score. Mottram and Donders found that the length of time in a coma was an influence on cluster types, with the low functioning clusters predicting longer comas. The low functioning cluster was also more likely to show signs of brain damage on neuroimaging findings.

Other studies have examined the validity of clusters that emerge from other clinical populations. For example, Waxman and Casey (2006) found that a five-cluster solution reliably represented a clinical population of children with learning disabilities and certain genetic disorders. Murji and colleagues (2003) used cluster analysis on individuals with HIV and found that derived clusters predicted performance on neuropsychological measures and clinical evaluations.

Standardization Samples

While cluster analytic studies of clinical samples can provide unique insights on how level and pattern of performance provide evidence of disease expression, they are limited if there is no available data on cluster profiles of normal controls for comparison. Therefore, researchers have also conducted cluster analysis studies on the standardization samples of selected IQ and neuropsychological batteries. For example, Donders (1996) looked at cluster subtypes formed from the WISC-III standardization sample. He did this on the premise that the WISC-III factor index scores are reliable determinants of levels and patterns of performance, based on a prior study by Glutting and colleagues (1994). He used the WISC-III index scores as attributes, and the 2,200 children from the standardization sample as objects. Consistent with the majority of the literature, he selected the Euclidean distance coefficient as the dissimilarity coefficient and Ward's method as the clustering method. He also ran a Complete Linkage method (CLINK), another agglomerative clustering method that is less likely to exhibit chaining, and the K-means partitional method to establish the clusters' stability. He found a five-cluster solution that differed on levels and patterns of performance. Specifically, three of the clusters represented overall levels of functioning that ranged from high to medium to low. The two other clusters were distinguished by the patterns of performance, in that one had a stronger processing speed index (PSI) score compared to the other index scores, and the other had a weaker PSI score.

Donders concludes from his findings that children can roughly be clustered into three levels of performance on the WISC-III or two patterns of performance that are primarily influenced by the PSI. He also suggests that as the analysis ignored age ranges, it is likely that the WISC-III accounts for age when measuring the PSI, as other research concludes that processing speed decreases with age in children (Mitchell, et al., 1990). One of the most interesting conclusions he makes is that the PS factor may be "contributing meaningful and unique information to the assessment process" (Donders, 1996; pp. 315).

In addition, Donders (2000) ran a cluster analysis on the CVLT-C standardization sample. Unlike the Wechsler measures, the CVLT-C does not provide factor index scores as easy and reliable sources for cluster attributes, so Donders selected five variables that best represented the five factors on the CVLT-C. These five factors are Attention Span, Learning Efficiency, Free Delayed Recall, Cued Delayed Recall, and Inaccurate Recall. The 920 children of the standardization sample ranged from 5-16 years of age, with equal proportions of boys and girls. As done in prior studies (e.g. Donders, 1996), both Ward's method and CLINK were selected as clustering methods. Furthermore, the K-means partitioning method was run to confirm the stability of the agglomerative clusters. Five clusters emerged from the standardization sample, three that differed by the level of performance and two by the pattern of performance. As with the WISC-III clusters, the three levels of performance clusters were differentiated as “high,” “average,” and “low” levels of memory functioning. A fourth cluster had low scores on the Attention Span subtest compared to other subtests, and a fifth cluster had low performance on all subtests except the Inaccurate Recall subtest, of which participants performed at the average level. Donders validated these clusters by comparing them across demographic variables. Children who came from lower socioeconomic backgrounds were typically in the low level cluster, while children whose parents completed college were typically in the high level cluster. Children with a history of learning disabilities or ADHD were more likely to be in the lower functioning clusters. Finally, age was not a contributing influence to differences among clusters, suggesting that the normative z-scores accurately correct differences that are due to age.

Donders and his colleagues (1996; 1997; 2000; 2006) have successfully found consistent cluster profiles that emerge on both the standardization samples and TBI samples of many neuropsychology measures. Generally, they found that five clusters emerge on standardization samples while four emerge on TBI samples. Three of the clusters vary on levels of performance, and the remaining cluster or clusters have different patterns of performance. Index scores and factors that represent processing speed seem to best account for profiles that vary across patterns, implying that processing speed is particularly sensitive to brain functioning.

Overview of the TOMAL

Reynolds and Bigler (1994) developed the TOMAL to fill a gap in memory batteries for children. Prior to its creation, the WRAML was perhaps the most cohesive and comprehensive memory battery designed for children and adolescents. Though the WRAML was a vast improvement over other measures of memory at the time, it was limited in the types of memory tasks it assessed (Reynolds & Bigler, 1997). Therefore, the TOMAL was designed partially to offer more breadth than previous psychometric measures of memory.

The TOMAL is composed of ten core subtests and four optional supplemental subtests. The core subtests are broken into verbal and nonverbal indexes, with five subtests per index. These two indexes can be combined into a composite memory index, which is an overall indicator of memory, similar to *g* on IQ measures. There is an additional index for delayed recall scores, as well as supplementary indexes for sequential recall, free recall, associative recall, learning, and attention /concentration. Finally, four empirically derived factor indexes are available (Reynolds & Bigler, 1996).

The TOMAL was unique for having an index for attention and concentration, which is a mitigating factor for performance on memory measures. Furthermore, the TOMAL has been noted for having nonverbal subtests that are difficult to encode verbally, promoting their validity as true measures of nonverbal memory performance (Reynolds & Bigler, 1997). This measure has shown to be a useful diagnostic tool for assessing learning disabilities (Howes et al., 1999), TBI (Lowther & Mayfield, 2004), ADHD (Morrison, 2006), and genetic disorders (Lajiness et al, 2005). Morrison's dissertation was particularly illuminating in confirming the TOMAL's discriminant validity in assessing ADHD, TBI, and controls. Reynolds and Bigler (1994) stated that the TOMAL moderately correlates with IQ measures, as is expected with a memory battery. Schmidt (2003) confirmed this with a TBI sample, noting that both the TOMAL and the WISC-III index scores were useful measures to distinguish mild, moderate, and severe pediatric TBI patients. Finally, Okura (2001) confirmed the construct validity of the TOMAL's verbal and nonverbal domains, with appropriate subtests correlating with the expected hemispheres of the brain.

Several studies have examined the factor structure of the TOMAL. Originally the authors (Reynolds and Bigler, 1996) proposed a two-factor solution for the TOMAL, splitting subtests into verbal and nonverbal domains, but after using both Varimax and Promax rotations, they determined that a four-factor solution was the optimal fit. These factors were highly stable across age groups and have been added as supplementary index scores. Follow up studies have been inconclusive on the stability of these factors across different clinical samples. For example, Alexander and Mayfield (2005) conducted a confirmatory factor analysis on a sample of 140 TBI patients and found that a two-factor model was a better fit than the four-factor model. However, rather than fitting into a "verbal" and "nonverbal" structure like the authors originally proposed, the two-factor model was divided into a general memory construct and smaller, attention span construct.

Lowe and colleagues (2002) investigated TOMAL gender differences using the standardization sample. Prior studies on gender differences in memory performance have produced mixed results. Some studies (Temple & Cornish, 1993) suggest that girls outperform boys on verbal memory tasks and others (Robinson et al., 1996) report that boys are superior on visual-spatial working memory tasks. Other studies (Ullman et al., 1997) report no significant relationship between gender and memory test performance. The participants of the TOMAL standardization sample consisted of 637 boys and 642 females. When factor analysis was conducted on each gender separately, the same four-factor solution emerged. Closer inspection indicated that a few individual subtests loaded on different factors, but that these differences were ultimately negligible. Further tests found that girls outperformed boys on two verbal tasks and boys outperformed girls on two visual tasks, but all such cases had small effect sizes that should not bias the test across genders. The authors conclude that whatever minor gender differences the TOMAL may have, those implications are inconsequential for clinical assessment.

In summary, the TOMAL has proven itself as a well-established memory battery for children and adolescents, and a useful diagnostic tool for a wide range of clinical disorders. Its widespread use and popularity have in part prompted the authors to increase the standardization sample and create the TOMAL-2.

Literature on the TOMAL-2

The TOMAL-2 (Reynolds & Voress, 2007) is the most recent revision of the TOMAL. The main addition to the TOMAL-2 is the increased age range, from 19 to 59 years. The subtests remain intact, with a few minor items changed. The authors state that more than 90% of the original items are unchanged. Only one new item is added to the Memory for Stories subtest, geared specifically for adults. Two of the TOMAL core subtests, Digits Forward and Visual Selective Reminding, were made supplementary on the TOMAL-2, shortening the average administration time to approximately 30 minutes. The reliability of the TOMAL-2 has been equivalent to the original TOMAL, and the same exploratory four-factor structure has emerged. Due to the test's recent publication, there are relatively few papers on the TOMAL-2 independent of the authors' studies. However, reviews for the TOMAL-2 have been favorable (Schmitt & Decker, 2008; Hartman, 2007), with particular praise towards its broad range of memory assessment techniques, its relatively cheap cost, and the abstract nature of the nonverbal subtests.

TOMAL and Cluster Analysis

There are a few studies that investigate the TOMAL clusters that emerge from clinical populations. Howes and colleagues (1999) ran two consecutive analyses on reading disabled children, using TOMAL subtest scores as attributes. Their first study compared the memory skills of children with reading disabilities with the memory skills of controls. The 135 children did not significantly differ on IQ scores. As expected, children with reading disabilities scored lower on the Composite Memory Index (CMI). The researchers' second study used cluster analysis to differentiate subtypes of the sample. The attributes were derived factor index scores with Eiganvalues over 1.0, and the objects were both the clinical and control samples. A six-cluster solution emerged from the data. The first cluster was comprised of control children who demonstrated strong to superior reading skills. The second cluster had children with very severe reading disabilities, which tied in with low performance across all memory index scores. The third cluster had mixed clinical and control children, and comprised of children who demonstrated weaker visual/spatial memory. The fourth cluster was also mixed and represented children with impaired verbal and auditory sequential memory. The fifth cluster only had two participants, both who exhibited severe nonverbal deficits. The sixth cluster had mostly control children who exhibited average performance.

The authors stated that the clusters reliably predict different cognitive profiles of children with reading disabilities. Specifically, the first and sixth clusters account for the majority of the controls, and can roughly be divided into “high functioning” and “average functioning” children. The second cluster had the most severe memory disabled children, and the third and fourth differentiated between verbal and visual/spatial memory deficits. Finally, the fifth cluster may represent a small population of children with severe nonverbal learning disabilities. The authors cautioned that the factor analysis that derived the index score attributes for the cluster analysis differed from the factors of the TOMAL standardization sample. However, as factor analyses studies on TBI TOMAL samples also differed (Alexander & Mayfield, 2005), it is reasonable to expect different factor structures to emerge from different clinical populations.

Fuller’s (2001) dissertation investigated the clusters that emerged from an ADHD sample. An initial factor analysis of the TOMAL yielded the same four-factor structure of the standardization sample: Complex Memory, Sequential Memory and Attention, Working Memory, and Spatial Memory. An eight-cluster solution emerged based on these four factor index scores. Two of these clusters were homogeneous for an ADHD diagnosis, while the other six had both ADHD and controls intermingled. Individual clusters differed on strengths and weaknesses on the TOMAL. These clusters significantly correlated with scores on other instruments used to assess attentional problems. Fuller concludes that children who have a comorbid diagnosis of ADHD and a reading disability exhibit the most severe deficits in memory.

Finally, Allen and colleagues (in preparation) investigated homogenous profiles of pediatric TBI and matched controls on the TOMAL. Clusters were compared across clinical, behavioral, and demographic variables to establish external validity. The researchers based their hypothesis on Mottram and Donders' (2006) paper on the CVLT-C. Mottram and Donders found that four clusters emerged in pediatric TBI patients, three that differed on levels of performance and one that had low performance on all variables except one that had an average level of performance. In the matched control sample, Allen and colleagues found a four-cluster solution that generally differed on level of performance. Three clusters distinguished high average, average, and low average scores while a fourth cluster had advanced scores on the NMI and VMI and superior ACI scores. For the TBI sample, Allen and colleagues found that two clusters emerged on their TBI sample that differed across levels of performance. However, three additional clusters identified impaired verbal, nonverbal, and attention performance respectively. These findings are consistent the other papers (Howes et al., 1999; Fuller, 2001) and suggest that the TOMAL is particularly adept at distinguishing impaired verbal, nonverbal, and attentional abilities. As the TOMAL has been previously praised for having nonverbal subtests that are completely independent of verbal abilities, the results are particularly promising in confirming the TOMAL's status as a broad measure of different memory domains.

External validation of the five-cluster solution showed that clusters did not differ due to gender, ethnicity, or method of TBI injury. However, the average level cluster was slightly older than the other clusters, and that the impaired cluster had significantly higher GCS. The average cluster had higher IQ scores on Wechsler tests. The findings in behavioral differences were of interest, as relatively few studies have compared clusters across behavioral variables. The impaired nonverbal cluster had the most self-reported clinical problems, while the average level cluster had the fewest problems across parent, and teacher report forms, yet high self reported problems, suggesting that even if a child has relatively few symptoms due to their TBI, they may still perceive themselves as damaged.

In summary, cluster analytic studies on the TOMAL have established that the TOMAL can adeptly distinguish clinical populations with TBI, ADHD, and reading disabilities. Furthermore, the TOMAL is promising as a strong measure that can identify weaknesses within verbal and nonverbal memory domains. However, there is a gap in the literature in that there have been no cluster studies on the original TOMAL standardization sample. Analyses on the standardization samples of other measures have shown that standardization samples have their own specific profiles compared to clinical samples. For example, the CVLT-C standardization sample (Donders, 1999) indicated that a five-cluster solution to be an optimal fit, while the CVLT-C pediatric sample (Mottram & Donders, 2006) yielded a four-cluster solution. Discrepancies in cluster solutions may indicate the presence of unique traits found within a clinical disorder. In turn, these traits may be identified endophenotypes that can serve as useful diagnostic markers for clinicians. Therefore, it is important to investigate the cluster solution of the TOMAL standardization sample, to see if TOMAL clusters reliably can distinguish verbal and nonverbal performances, or to see if a new cluster solution may emerge. As the TOMAL-2 contains two fewer core subtests than the TOMAL, it is important to investigate the different cluster solutions that may emerge.

Hypotheses

We proposed that consistent with the prior research (Allen et al., in preparation), four clusters will emerge from the TOMAL standardization sample that will largely differ across level of performance. One group will exhibit high average memory performance, a second will indicate average performance, and a third will have low average performance. Finally, a fourth cluster will have advanced memory performance with a superior ACI score.

We also hypothesize that there will be differences in the cluster solutions for younger (age 5-8), middle (9-11), and older (ages 12-19) children given the developmental changes in cognitive abilities. The cutoff scores of 9 and 12 were decided based on reviews of memory development in children (Gathercole, 1998; 1999). We cannot make specific predictions regarding these differences, as there is no basis in the literature to do so. Based on the available literature, we do point out that attention/concentration abilities (Swanson, 1999) and verbal/nonverbal memory differentiation (Gathercole, 1998) may be areas where the younger and older children are differentiated with regard to the cluster analysis results.

CHAPTER 3

METHODOLOGY

Participants

Participants were 1,121 children and adolescents who were selected from the 1,342 individuals from the TOMAL standardization sample. The standardization sample was comprised of children who were stratified by gender, ethnicity, socioeconomic status, geographic region, and urban/rural resident to match the normal population of the United States based on census data from 2000. Children who had incomplete subtest profiles and children with learning disabilities (LD) were excluded from this study; all other children from the standardization sample were retained. Our sample was between the ages of 4 years, 6 months, and 19 years, 10 months. Ages were rounded into the following classifications: 67 children were age five, 78 were age six, 124 were age seven, 92 were age eight, 89 were age nine, 143 were age ten, 145 were age eleven, 97 were age twelve, 68 were age thirteen, 59 were age fourteen, 38 were age fifteen, 28 were age sixteen, 40 were age seventeen, 34 were age eighteen, and 19 were age nineteen.

Males comprised 48.1% of the sample (n = 539), females comprised 48.7% (n = 546), and 36 participants had unidentified gender. Ethnicity was divided into Anglo/European (70.0%), African-American (13.2%), Hispanic (8.9%), Oriental/Pacific Islander (2.1%), and Native American (2.4%) categories. With regard to race, 83.8% identified as Caucasian, 13.0% identified as African-American, and 3.52 identified as Other. Geographic regions included the following: Northeast (13.8%), North Central (25.1%), South (36.7%), and West (24.4%). Children were tested in the states of California, Colorado, Florida, Georgia, Maine, Maryland, Massachusetts, Michigan, Minnesota, New Hampshire, New York, North Carolina, Ohio, Pennsylvania, Texas, Utah, and Washington. Teams of trained examiners went to elementary and high schools in all regions to administer the test to volunteers.

TOMAL and TOMAL-2

The TOMAL is standardized for use with children and adolescents. The TOMAL has a core battery with ten subtests which are further divided into the Verbal Memory Index (VMI) and the Nonverbal Memory Index (NMI), each which contain five of the core subtests. All subtests together form the Composite Memory Index (CMI). Two of the verbal subtests and two of the nonverbal subtests can be administered a second time 30 minutes after the initial administration to form a Delayed Recall Index (DRI). In addition to these subtests, four supplementary subtests are provided. Three of the subtests are verbal and one is nonverbal. The authors state that one of the supplementary subtests can replace a core subtest if a core subtest is unavailable. A combination of both core and supplementary subtests form several additional supplementary indexes: Sequential Recall Index (SRI), Free Recall Index (FRI), Associative Recall Index (ARI), Learning Index (LI), and Attention/Concentration Index (ACI). Each subtest has a mean of 10 and a standard deviation of 3. Summed scaled subtest scores are calculated into index scores which have a mean of 100 and a standard deviation of 15. In the case of missing subtests, prorating methods are available to calculate representative index scores. Percentiles are provided alongside standard scores.

The Memory for Stories subtest (MFS) is a verbal subtest that measures semantic and sequential recall. This subtest consists of five stories that are read outloud by the examiner. Depending on the child's age, the examiner will start on Story 1, Story 2, or Story 3 and the examiner always discontinues after three stories are completed. The examiner also discontinues if the child scores a 0 on any of the stories. The examiner begins by telling the child that he will read her a short story. After the story is read, the examiner prompts the child to tell the story back as best as she can. The child will relate back as many details as she can remember. The order in which she relays the details is unimportant but semantic details (i.e. names, dates) are worth points. Upon completing MFS, the examiner can start a timer so that in 30 minutes, a delayed MFS can be administered. With the delayed MFS, the examiner prompts the child to relay as many details about the three stories as possible, without reading the stories a second time.

The Facial Memory (FM) subtest requires the child to recognize several black and white photos of faces. These faces are of different ethnicities, gender, and age ranges. Facial memory is a unique form of nonverbal memory that has its own distinct forms of aphasia (i.e. prospagnosia). A stimulus book has several pages of faces. Each item is administered over two pages. On the first page are one or more faces. The child is given some plastic chips and prompted to study the face or faces. After a preset number of seconds the examiner turns the page to a page with several faces, one or some of which were the original target face or faces. The child is then asked to place the chips on the target faces from the previous page. The child is always given the exact same number of chips as target faces for each item. Earlier items require the child to view the target for 5 seconds. Items with more faces increase the observation time to 10, 15, and 20 seconds to allow the child to process the additional target faces. One point is awarded for each correct face that is identified.

The Word Selective Reminding subtest (WSR) is a test of learning and immediate verbal recall. Learning tests are distinct on the TOMAL because the examiner will remind the child of any target words or stimuli that were left out of a previous trial. Learning rates are identified by how quickly a child masters a word or stimulus list. In the WSR, the examiner reads off a list of words to the child. Afterword, the child is prompted to say back as many of the words as she can remember. The order of recall does not affect scores, but is marked down by the examiner. After each trial the examiner reminds the child of which words she forgot, and then he reads the word list again. When a child remembers every word perfectly on two consecutive trials, the subtest ends and the child receives full credit for any subsequent trials she did not have to complete. The subtest also ends after eight trials, regardless of memory mastery.

The Visual Selective Reminding (VSR) subtest is similar to the WSR with a visual rather than verbal modality. Therefore, the VSR can be considered a test of visual learning and immediate recall. The examiner displays a test board with several dots, each containing a single digit, scattered across the board. He prompts the participant by touching the dots in a predetermined order in a steady rhythm. Immediately after finishing, the participant is prompted to touch the dots. The order the participant touches the dots does not need to match the examiner's predetermined order, rather the participant simply has to touch all the dots the examiner touched.

Like all learning tests, this subtest is comprised of a series of trials. At the end of each trial, the examiner corrects any errors the participant made in order to evaluate learning rates. This subtest is discontinued either after eight trials, or when the participant gets a perfect score on two consecutive trials.

Object Recall (OR) is the fifth core subtest administered and is unique in tasking the child to transfer information from one modality (visual) into another (verbal). This sort of information processing may be impaired by learning disabilities. The examiner presents a picture book to the child. The examiner flips through the picture book, naming each object outloud to the child. Upon completion, the examiner prompts the child to remember as many of the objects as possible and to respond verbally. Although OR has visual stimuli and verbal responses, it is considered a verbal subtest and contributes solely to the VMI.

Abstract Visual Memory (AVM) has a similar format as FM, although instead of faces the child is shown various abstract images. There is also only one target stimulus per item. The examiner shows the participant a page with an abstract design, and then flips to a second page where several abstract designs, one being the original target design, are displayed. The child has five seconds to respond to the correct design. This subtest is considered a measure of pure nonverbal memory, as the child does not need to verbally identify the stimuli.

Digits Forward (DF) is a measure of rote memory and is considered a verbal subtest on the TOMAL (but not the TOMAL-2). Its structure is almost identical to that of the Digit Span subtest from the Wechsler scales. The examiner reads a list of predetermined numbers and prompts the child to repeat the numbers in the exact order.

This test can be considered a measure of attention and concentration as well as rote memory, as children with attention difficulties may have trouble encoding the stimuli.

The Visual Sequential Memory (VSM) subtest is a measure of abstract nonverbal sequential memory. The examiner presents the child with a picture book in which a series of designs are laid out. The designs are purposefully abstract and meaningless geometric shapes. The shapes are in a specific order on the first page. The child is allowed five seconds to examine the page, and then the examiner turns to the second page and prompts the child to point to the pictures in the order they were previously displayed. The child only receives a point if she gets the entire sequence correct.

Paired Recall (PR) is a verbal learning subtest in which the child must pair a prompted word with an associated word. The examiner reads off a list of paired words to the child. Easier items pair words with easy semantic connections while more difficult items have pairs with fewer semantic connections. After reading off a list of paired words, the examiner then prompts the child by saying the first word of each pair. If a child is incorrect, he is corrected with the right answer. Subsequent trials are composed of the same word pairs, but the examiner reads them in a different order.

Memory for Location (MFL) is a nonverbal subtest in which the child is tasked to recall the locations of dots on a grid. The examiner presents the child with a picture where several large dots are randomly arranged. The child is allowed to study the picture for five seconds and then the page is turned. The second page displays a grid. The child is prompted to point to the grid boxes in which dots would appear from the previous page. Each correct answer is worth one raw point. The child is not corrected for wrong answers after the initial trial period.

The Manual Imitation (MI) supplementary subtest is unique in that it assesses motor skills as well as immediate memory. Both the examiner and the child must remove all jewelry from their right hand. The examiner then shows the child four basic hand motions: palm down, fist, palm up, and side. The child is then instructed to watch the examiner make these various hand motion at the rate of one per second. One raw point is awarded for each correct motion in the sequence. After the initial trial period, the child is not corrected for wrong hand motions.

Letters Forward (LF) is a supplementary verbal subtest that is similar in modality with the DF subtest. Instead of reciting digits, the examiner prompts the child with random letters. As item difficulty increases, the length of each string of letters increases by one.

Digits Backward (DB) and Letters Backward (LB) are the other two supplementary subtests. In both subtests, the examiner prompts the child with a string of numbers or letters and instructs the child to recite them in reverse order. Some research has suggested that the backwards span tasks tap into working memory rather than pure rote memory (Ramsay & Reynolds, 1995).

The subtest makeup of index scores on the TOMAL and TOMAL-2 slightly differ. Five subtests make up the VMI on the TOMAL: MFS, WSR, OR, DF, and PR but only four subtests are used on the TOMAL-2: MFS, WSR, OR, and PR. The DF subtest was made supplementary on the TOMAL-2. Five subtests also make up the NMI on the TOMAL: FM, VSR, AVM, VSM, and MFL and only four subtests are used on the TOMAL-2: FM, AVM, VSM, and MFL. Like the DF subtest, the VSR subtest was consigned to a supplementary status. This reduces the number of core subtests on the TOMAL-2 from ten to eight and is the main reason the TOMAL-2 has a shorter administration time.

Supplementary indexes are for the most part identical on the TOMAL and TOMAL-2 with some exceptions; for example, the FRI is composed of the FM, OR, AVM, and MFL subtests on the TOMAL and only composed of the FM, AVM, and MFL subtests on the TOMAL-2. Other supplementary indexes are made up of the following subtests for both TOMAL versions: the ACI is composed of the DF, LF, MI, DB, and LB subtests; the SRI is composed of the VSM, DF, LF, and MI subtests; the ARI is composed of the MFS and PR subtests; and the LI is composed of the WSR, OR, PR, and VSR subtests. Finally, the DRI is composed of the MFSD, FMD, WSRD, and VSRD on the TOMAL. On the TOMAL-2, it is renamed the Verbal Delayed Recall Index (VDRI) and is composed only of the MFSD and WSRD. See Table 1 for a complete breakdown of subtest composition.

Reliability

The authors determined internal consistency reliability using Cronbach's coefficient alpha (Cronbach, 1951). Cronbach's alpha measures how a group of variables consistently measure a single construct. It increases when intercorrelations among items increase and indicates that there is less random error in the scores. The authors determined that the internal consistencies of the subtests are very high, typically exceeding coefficients of .90. Except for some of the delayed recall subtests, no coefficients are below .74. As reliabilities in the .70s are still considered quite high by most standards, the TOMAL's internal consistency is considered to be excellent. Internal consistencies of index scores were also high with most coefficients around .90 and core indexes above .95. As expected due to very similar item content, the TOMAL-2 also possesses very high internal consistency.

Stability was established with test-retest reliability. A measure with high test-retest reliability is expected to have no significant changes in performance across subsequent administrations on the same individuals. It measures the stability of individual performance, the measured trait, and errors in assessment due to time sampling. The authors assessed 35 children twice at an average six week interval and found that reliability was typically in the .80s. The authors add that differences across subtests and index scores were less .2 standard deviations apart, confirming strong reliability. For the TOMAL-2 the authors conducted reliability tests on both children and adults and found results akin to that of the TOMAL.

Validity

Content validity was established throughout the item development phase, in which the authors used experts, pilot studies, focus groups, memory theory, and empirical findings to form items with relevant content to the constructs they were trying to measure. Most of the subtests were devised a priori to fit either a verbal or nonverbal category in order to match the theory of differentiating memory processes. Tryout phases allowed the examiners, examinees, and expert witnesses to provide feedback on item content. The item development phase took almost three years to complete. Once completed, norming and standardization processes followed.

The authors also ran factor analyses using the correlation matrices of the 14 subtests to identify stable constructs. A confirmatory analysis examined the fit of the verbal and nonverbal subtests that formed the VMI and NMI. A second exploratory analysis was also examined. The authors elected to use an oblique rotation with the Promax procedure. The two-factor solution found that most of the subtests loaded onto a general memory factor and the remaining subtests on a smaller, attention/concentration factor. The second factor was named this because the DF, LF, DB, LB, and MI subtests had the strongest loadings on this factor and these subtests form the ACI. The exploratory analysis found a four-factor solution comprised of a general memory factor, a sequential recall and attention factor, a backwards recall factor, and a spatial memory factor. All factor solutions were consistent across age ranges and emerged again on a subsequent analysis of the TOMAL-2.

Finally, criterion-related validity studies were conducted. The authors found that the TOMAL moderately correlated with IQ and achievement tests. More specifically, the VMI had better correlations with achievement tests than the NMI, with ranges from .28 to .46, while the NMI had ranges from .07 to .25. In regard to IQ measures, index scores correlated higher with expected IQ indexes of similar modality, i.e. verbal indexes correlated highly together as did nonverbal indexes. The authors conclude that these significant yet moderate correlations indicate that the TOMAL is related to measures of intelligence and achievement but at rates lower than typical between measures of intelligence or achievement. With a shared variance ranging from 16% to 36%, the TOMAL adds its own contribution to a comprehensive assessment. The TOMAL-2 expounded on this by correlating the TOMAL to the

WRAML-2 and finding high correlation coefficients that exceeded correlations with measures of intelligence and achievement. This adds further evidence that the memory measures are capturing constructs that are separate (but linked) to intelligence constructs, i.e. memory and learning.

Reynolds and Bigler (1994) compared TOMAL index score performance of 41 children with learning disabilities (LD) to the standardization means and found that the LD children performed almost a full standard deviation below the SS on the ACI and SRI, but performed near the population mean of the ARI. Subsequent studies by outside researchers have found that the TOMAL is also sensitive to TBI (Lowther & Mayfield, 2004), reading disabilities (Howes, Bigler, Lawson, & Burlingham, 1999), ADHD (Thaler, Allen, McMurray, & Mayfield, in press), and genetic disorders (Lajiness et al., 2005).

Data Analysis

A series of cluster analyses was run on the 1,121 individuals selected from the TOMAL standardization sample. The descriptive and exploratory nature of cluster analysis necessitates that the experimenters have a strong theoretical rationale before analyzing the data. Therefore, several factors must be taken into account before running a cluster analysis such as choosing the objects, the attributes, the resemblance coefficients, the clustering methods, and the final number of clusters. In addition, all analyses were divided into "younger children," "middle children," "older children," and "combined." Cluster analyses was used to determine a potential natural cutoff point for these different age groups by forcing a three-cluster solution and examining the mean ages of all groups. If no such cutoff was identified, the cutoff points were selected as ages 9 and 12, as previous studies suggest this is a critical point in memory development (i.e. Gathercole, 1998; 1999).

In our study the objects were the participants, as typically is the in the analyses of psychometric measures. However, the selection of attributes can vary. For example, in prior TOMAL cluster analysis studies researchers selected attributes as both empirically derived factors (Howes et al., 1999) or the TOMAL subtests (Allen et al., in preparation). We have elected to use standardized subtests as attributes for this study, as they provide a more relevant clinical portrayal of the domains assessed by the TOMAL.

We used only one resemblance coefficient across all analyses: the squared Euclidean distance. This was decided due to the universal use of this resemblance coefficient across all cluster analysis studies of neuropsychological measures (i.e. Donders, 1996; Donders & Warchausky, 1997; 1999; Goldstein, Allen, & Seaton, 1998; Mottram & Donders, 2006; Seaton et al, 1999). The Euclidean distance coefficient is a dissimilarity coefficient in which the smaller the value, the more similar the two objects (Romesburg, 1994). This coefficient calculates the distance between two objects based on all their attributes using a form of the Pythagorean theorem. When there are only two attributes, this coefficient may be best envisioned as the length of the hypotenuse between two points of a right triangle As attributes increase it becomes more difficult to envision the distance coefficient but the mathematics remain the same; the close two objects are in Euclidean space, the more they are similar across the various attributes.

Unlike the resemblance coefficient, psychometric researchers have used different clustering methods for their analyses. This study examined the two-step method and Ward's minimum variance method. The two-step method empirically creates many small sub-clusters, and then fits these sub-clusters into the specificed number of clusters (Bachler, Wenzig, & Vogler, 2004). Ward's method forms larger clusters based on whichever merger will have the smallest increase in variance, which is determined by a sum of squares formula (Ward, 1963). The two-step method was used as an exploratory method while Ward's method will specified number of clusters *a priori* to analysis.

A review of previous TOMAL cluster studies suggested that the TOMAL may be best interpreted as a five cluster solution in children with TBI but four in the standardization sample (Allen et al., in preparation). To maximize our use of the data we ran analyses of four, five, and six cluster solutions using Ward's method. Along with looking at the average attributes for each cluster, clusters were visually plotted on a line graph across the VMI, NMI, and ACI. These three indexes were specifically selected for clinical relevance, as they assess three separate constructs (verbal memory, nonverbal memory, and attention/concentration) and have minimal subtest overlap (only the DF subtest is shared by the VMI and ACI). Further, subtests that form the DRI were not included in the analysis due to their restricted range, as discussed by the TOMAL's authors (Reynolds & Bigler, 1994; 1996).

In addition to plotting clusters across the VMI, NMI, and ACI, we plotted the clusters across TOMAL factor scores. Although limited for clinical use, an examination of factor patterns may provide insight on separate constructs of memory, as factor scores have the advantage of being completely orthogonal. See Table 2 for the four-factor solution of the TOMAL standardization sample. As reported by Reynolds and Bigler (1994), the first factor is composed of a variety of general verbal and nonverbal memory tasks and captures overall complex memory skills. The second factor is composed of the DF, LF, VSM, and MI subtests and loads subtests that emphasize forward sequential recall and attention. The third factor has the two backwards span tasks. Backwards span tasks may require more working memory than forward span tasks (Ramsey & Reynolds, 1995), explaining the separate loading between the forward and backward tasks. The fourth factor is composed of subtests that tap into spatial memory. Factor scores for our sample were calculated by summing the subtests contained within each factor, taking the average, and then standardizing these scores. Cluster solutions were then plotted across the score to see how profiles vary among the four factors.

Table 3 has a breakdown of the different age groups, methods, plotting patterns, and cluster solutions that will be examined for this study. As seen in the table, multiple clustering solutions were investigated in order to determine the best possible solutions. First, a series of cluster analyses looked at combined scaled scores and identified the best solution using both Ward's Minimum Variance and the two-step solution. Second, an exploratory cluster analysis using raw scores determined the best cutoff age to split younger and older children (if no cutoff age was identified, we set at ages 9 and 12). Finally, cluster solutions were run on standards scores for the younger, older, and combined groups using both the Ward's method and two-step method, then plotted across the VMI, NMI, ACI, and TOMAL factor scores. The entire standardization sample was analyzed first in order to determine which clusters emerged when age groups were combined. This allowed us to see which clusters in the three age groups reflected the clusters from the entire standardization sample, and which clusters were unique, and thereby provide insights about the nature of memory in learning in their respective age groups.

As each cluster solution was identified, its internal validity and stability were determined. Internal validity was established by graphing the clusters in discriminant function space. Stability was determined by running a second-stage K-means iterative partitioning cluster analysis with subtest/factor score means specified as starting points for each cluster centroid. The k-means method is nonhierarchical in which predetermined centroids or "seeds" are established as the center of a predetermined number of clusters (McIntyre & Blashfield, 1980). Other objects are then assigned to a cluster based on their distance to each seed. A stable cluster solution would theoretically have similar cluster membership with both the hierarchical and iterative methods. Cohen's Kappa and discriminant function analyses (DFA) was used to establish the level of agreement between the two cluster stages. Finally, Beale's F-statistic (1969) was used to determine which cluster solution was the most parsimonious compared to other cluster solutions. The combination of using Cohen's Kappa and DFA for stability and Beale's F-statistic for parsimony helped us identify which clustering solutions were the most optimal for the three age groupings we looked at in the TOMAL.

External validity of optimal cluster solutions were assessed by comparing our final cluster solutions on variables not included in to cluster analyses, including age and gender. One way ANOVAs compared the groups on age and chi-square analysis compared across gender to establish external validity.

CHAPTER 4

RESULTS

Cluster Analyses of the Entire Sample

A four, five, and six-cluster analysis was run using Ward's method and the two-step method on the entire sample using scaled subtest scores as attributes. See Table 3 for the average subtest scores of the four-cluster solution using Ward's method. See Table 4 for the average subtest scores of the five-cluster solution using Ward's method. See Table 5 for the average subtest scores of the six-cluster solution using Ward's method. These solutions were plotted along the VMI, NMI, and ACI index scores as seen on Figures 2, 3, and 4.

As seen in Figure 2, clusters are differentiated by both level and pattern of performance. For the four cluster solution, an identified low average cluster (C1) emerged in which performance across all three indices was less than 90. A second high average cluster (C4) emerged with higher VMI and NMI index scores and an even higher ACI index score that averaged over 125. A third cluster (C2) performed approximately average with a slight dip in the NMI. The fourth cluster (C3) had an elevated NMI, suggesting individuals in this cluster had stronger nonverbal capabilities.

The five cluster solution had the low average (C1), high average (C5), and strong nonverbal (C2) clusters as well. However, the average cluster split into two additional clusters. One cluster (C3) remained average while the new cluster (C4) had a slightly elevated VMI and a higher ACI that exceeded 110. The six cluster solution retained all previous clusters. However, the high average cluster was split into two more clusters: one that was superior (C6) with VMI and NMI over 120 and ACI over 130, and one that remained high average (C5). Both these clusters exhibited the same pattern of performance with strong VMI and NMI capitulated by an even stronger ACI.

Next, cluster solutions were plotted across the standardization sample's four factor scores. Factor scores were calculated by averaging the standard scores of the subtests that make up each factor. For example, the Backwards Recall/Working Memory factor score was derived by summing the two subtests that make up this factor (Letters Backwards and Digits Backwards) and dividing that sum by two. The Backwards Recall/Working Memory factor seems to play a role in elevating the high average cluster (C4) with an average subtest score approximating 14.5. The low average cluster (C1) performed low on all subtests, while the other two clusters were similar on most of the factors, except that one cluster scored higher on the spatial memory factor (C3; over 12 points) and the other scored lower (C2; less than 9 points). When a fifth cluster was added, the new cluster appeared to have strong attention and working memory scores but unlike the high average cluster, this cluster did not score particularly high on the general memory factor. The addition of a sixth cluster followed a similar pattern, with the new cluster exhibiting strong scores on the working memory factor as well as somewhat elevated general memory and attention factor scores. See Figure 5 for the factor profiles of the five-cluster solution.

Cluster stability of the solutions was the examined first using DFA. DFA correctly classified 83.7% of the cases into the correct cluster of the four-cluster solution, 79.6% of the five-cluster solution, and 79.2% of the six-cluster solution.

Charts of the discriminant function space are seen in Figures 6, 7, and 8. Stability was next evaluated using an alternative clustering method for reliability: the K-means classification process. Centroids were specified as the means for each TOMAL subtest in standard score format per cluster specified by Ward's method. Cross-tabulation compared the emerged clusters with the original hierarchical solution and Cohen's Kappa was used as a measure of stability. Tables 7, 8, and 9 display the four, five, and six cluster solutions' classifications respectively. The Kappa for the four-cluster solution was .69; for the five-cluster solution it was .67, and for the six-cluster solution it was .66. These preliminary results suggest a moderate degree of stability between the two solutions. Although the four-cluster solution has a slightly better classification rate than the 5 and 6 cluster solutions, the three solutions appear to have comparable stability.

A second cluster analysis was then run using the two-step clustering method. Unlike Ward's method, the two-step clustering method automatically selected the optimal number of clusters. The two-step clustering method only extracted two clusters that represented high functioning and low functioning participants. As these two clusters lacked any discernable pattern differences, this two-cluster solution was not further analyzed. However, as an alternative method the TOMAL raw subtest scores were selected as attributes for a second analysis. When raw scores were used, the clustering method determined that a three-cluster solution was optimal for the standardization sample. Table 10 displays the mean subtest scores in scaled form, Figure 9 displays the VMI, NMI, and ACI performance across the three clusters. As seen in Figure 9, these three clusters mostly vary on level of performance, with high, middle, and low average clusters. Of note, the high average cluster fits the same profile as the clusters derived from Ward's method (i.e. high VMI and NMI with an even higher ACI). However, the spread of scores is somewhat constrained; the low average and average group do not differ greatly and in fact, converge across the ACI (indicating that the low average cluster may be better termed as a second average cluster). When plotted across factors as seen in Figure 12, the high average cluster still remained distinct (C3), although the other two clusters were much closer together than on other cluster profiles.

Like with the other clusters, this cluster's stability was assessed using DFA and Cohen's Kappa. The DFA correctly classified clusters with raw subtest scores 94.2% of the time (see Figure 10 for a plot of the discriminant function space). The clustering solution was reevaluated using the K-means iterative process, specifying mean raw subtest scores for each clusters as centroids. Cross-tabulation classification rates are on Table 11, with Cohen's Kappa at .85. From this it appears that the two-step, three-cluster solution has a better classification rate compared to the four, five, or six cluster solutions derived through Ward's methods. As raw scores were used to calculate this profile, it is likely that the two-step profile differs greatly on age. A one-way ANOVA was used to confirm that this solution's clusters significantly differ by age, $F(2, 1118) = 527.68$, $p < .001$. Descriptives indicated that the low functioning cluster had an average age of 7.1 ($SD = 1.8$), the middle functioning cluster had an average age of 11.1 ($SD = 2.7$), and the high functioning cluster had an average age of 13.5 ($SD = 2.9$). Therefore, despite the high stability of this cluster solution, the interpretation of cluster profiles is limited by the effects of age.

To see if a three-cluster method using Ward's method would have a similar classification rate, a final Ward's method clustering solution was derived with three clusters specified. The clusters here resemble those found in the two-step method, although with a greater spread of scores; the high average group (C3) performed approximately a standard deviation above the mean while the low average group (C1) performed about one standard deviation below the mean. When plotted across factor scores, the three clusters distinctly varied on level of performance, and the high average cluster (C3) displayed an even higher working memory factor score. DFA correctly classified the clusters 85.7% of the time and Cohen's Kappa with a separate K-means clustering solution was .69.

Finally, the three, four, five, and six cluster solutions were evaluated using Beale's *F* statistic (Beale, 1969). This statistic evaluates clusters for parismony by comparing two clusters' sums of the squared Euclidean distances from their respective centroids against the critical value of the *F* distribution. Clusters that account for significantly more variance are considered superior and are retained. As three simultaneous tests were assessed, a Bonferroni correction set the p-corrected value at .017 (.05/3). A comparison of the three and four cluster solutions found that the four cluster solution did not account for significantly more variance than the three cluster solution, $F(14, 15638) = 1.80$, $p = .03$. A comparison of the four and five cluster solutions found that the five cluster solution did account for significantly more variance than the four cluster solution, $F(14, 15624) = 2.62$, $p < .001$. Finally, a comparison of the five and six cluster solutions found that the six cluster solution did not account for significantly more variance than the five cluster solution, $F(14, 15610)$

= 1.80, p = .03. From this it appears that the five cluster solution is the optimal fit for the entire TOMAL standardization sample.

A chi-square analysis found no significant differences in the proportions of boys to girls among the five cluster solution, $\chi^2(8) = 11.0$, $p = .20$. However, a one-way ANOVA did find significant differences due to age, $F(4, 1116) = 7.42$, $p < .01$. Bonferroni analyses found that the high average cluster (C5) was significantly older than the low average cluster (C1), the average cluster (C2), and the nonverbal cluster (C3). However, this cluster was not significantly older than the attention cluster (C4). In addition, the attention cluster was significantly older than the average cluster. Table 12 has a breakdown of the age differences among the clusters.

Prevalence rates for these five clusters differed. The low average cluster captured the most children and represented approximately 28% of the cluster. Conversely, the high average cluster only captured 13% of the population while the average cluster captured 19%. The nonverbal cluster was more common with 25% of the sample, while the attention cluster represented 15%.

Differences by Age

In order to provide a visual aid in determining an optimal cutoff age, TOMAL subtests were standardized into z-scores and then plotted across the different age groups as seen on Figure 11. As expected, most raw subtest scores steadily increased with age. The rate that scores increased was variable, although a look at the point between the ages of 8 and 9 indicate a jump for at least some subtests, such as WSR, DF, LF, and VSM. A second increase of raw scores appears to occur between the ages of 11 and 12, as seen with subtests including the LF, MI, MFS, and LB. This suggests in part that cognitive changes that occur at these critical ages may lead to larger increases in memory ability. A look at age groups older than 12 shows a more steady increase of subtest performance, indicating that memory develops somewhat slower after the initial years. From a purely visual perspective, it appeared that age 8 and age 11 may be adequate cutoff points to differentiate the cluster groups.

As already analyzed, the two-step exploratory cluster using raw scores extracted three clusters that differed primarily on age, with a young, middle, and older group emerging. A second cluster analysis was run specifying these three clusters, using Ward's method with raw subtest scores as attributes. As anticipated, the three clusters significantly differed on age, $F(2, 1118) = 311.92$, $p < .01$. Post-hoc Scheffe tests confirmed that all three clusters significantly differed from each other by age. When mean ages were inspected, it was found that one cluster had a mean age of 8.2 ($SD = 2.7$; N = 476), a second cluster had a mean age of 11.5 ($SD = 3.0$, N = 394), and a third cluster had a mean age of 13.3 ($SD = 2.9$, N = 251). With the exception of the younger cluster, these mean ages matched the two-step cluster mean ages. The two-step method found that the younger cluster had a mean age of 7.1, although this only differs from Ward's method by one year. Overall, as the findings coincide closely with Figure 16 and the two-step analysis and it was deemed that three age groups should be separately analyzed, one ranging from age 5-8, a second ranging from 9-11, and a third ranging from 12-19. All subsequently cluster analyses were run on each of these three age groups.

Young Age Group (5-8 Years)

As with the entire standardization sample, a four, five, and six cluster analysis were run on the young age group, using Ward's method and standard scores as attributes. A second exploratory analysis was then run using the two-step method and raw scores as attributes. When using Ward's method, the five and six cluster solution each produced one cluster with only six participants. When inspected, these six participant clusters seemed to capture children who were exceptionally gifted, with scores that were an average of two standard deviations over the population mean. However, due to the small size of these clusters, they were deemed clinically irrelevant for the purposes of the study and removed. Therefore, only the four-cluster solution was initially analyzed. See Table 13 for means and standard deviations of the subtest scores for the four-cluster solution. In addition, the four cluster solution was plotted across index scores on Figure 12 and was plotted across factor scores on Figure 13.

When looking at the plotted index scores, it is immediately apparent that compared to the entire standardization sample, the young group had much less variation on pattern of performance and more levels of performance. In the four cluster solution, low average (C1), average below the mean (C2), average above the mean (C3), and high average (C4) clusters emerged. The two average clusters were close together, but differed more on their VCI and ACI indexes. The high average cluster had a slight elevation on the ACI compared to the VMI and NMI, though not as pronounced as the high average cluster for the entire standardization sample. For the factor scores, the same pattern emerges – young children have steady levels of performance that can be roughly categorized as low, low average, high average, and high.

DFA correctly classified 87.3% of the cases, a level comparable to the entire standardization sample. When the K-means iterative analysis was run, the two cluster solutions had a Kappa of .67, an acceptable agreement rate. See Table 14 for agreement rates between the two cluster solutions.

Next, the two-step method ran an exploratory cluster analysis on the 5-8 years group. This method found a two-cluster solution for this age group. An inspection of the clusters indicates that one cluster had 93.6% of the cases (n = 338). The other cluster (n = 23) captured children who scored exceptionally high on the TOMAL, similar to the extra clusters that emerged on the five and six-cluster solutions. Given the imbalance of children among these two clusters, the two-step cluster solution was not further analyzed.

Due to the lack of meaningful findings for the five and six cluster solutions and the two-step method, a final three-cluster solution was analyzed using Ward's method, as it may be that the three-cluster solution is indeed a better representation of the younger children than the four-cluster solution. Figure 14 plots this solution on TOMAL index scores. As expected, the three-cluster solution separated the young children into high average (C3), average (C2), and low average (C1) clusters with little variation on pattern of performance. DFA for the three-cluster solution correctly classified 86.4% of the cases and the kappa agreement with the K-means method was .73. See Table 15 for the cross-tabulation agreement rates between the Wards and K-means methods.

To determine which of the two clusters the better solution is, Beale's F statistic was calculated to determine whether the four-cluster solution accounted for significantly more variance than the three-cluster solution. Beale's F did find that the four-cluster solution accounted for significantly more variance than the three-cluster solution, $F(14, 4998) = 2.15$, $p < .01$, indicating that the four-cluster solution is optimal in younger children. A comparison across gender on the four-cluster solution found no significant differences among clusters, $\chi^2(3) = 2.42$, $p = .49$, and a one-way ANOVA found no significant differences in age due to cluster membership, $F(3, 357) = 2.51$, $p > .05$. The prevalence rates for clusters were 22% for the low average cluster, 20% for the average cluster that was below the mean, 41% for the average cluster that was above the mean, and 18% for the high average cluster.

Middle Age Group (9-11 Years)

Again, a cluster analysis was conducted for four, five, and six solutions using Ward's method and TOMAL subtest scores as attributes. Unlike the younger group, all three cluster solutions had several participants in each cluster, ranging from 25 up to 171. Therefore, all cluster solutions were inspected. See Tables 16, 17, and 18 for the means and standard deviations of the standardized subtest scores per cluster. An initial look at the data reveals one cluster that stands out as having exceptionally high Memory for Location scores, ranging around 17 to 18. This is unexpected given this cluster's relatively average performance on the other subtests. The clusters were further investigated across TOMAL index scores as seen on Figures 15, 16, and 17. As seen in Figure 15, the four-cluster solution is differentiated both by level and pattern of performance with a low average cluster (C1), an average cluster (C2), an advanced cluster (C4), and an average cluster with particularly strong nonverbal scores (C3). This cluster has the high Memory for Location scores, along with elevated scores in other nonverbal subtests. In the five-cluster solution the average cluster is split into a cluster with higher verbal and attention scores and average nonverbal scores (C2) and a cluster with higher verbal and nonverbal scores and average attention scores (C3), essentially removing an average performing cluster completely. The six-cluster solution sees a reemergence of an average cluster (C2) which splits from the low average cluster. Thus, in contrast to the younger age group, the middle age group has clusters that vary greatly across the three index scores.

Next, the three cluster solutions were plotted across TOMAL factor scores. See Figure 18 for the 5-cluster solution's plot across factor scores. When plotted across factor scores, the four-cluster solution has an advanced (C4), average (C2), and low average (C1) clusters as well as a cluster with average performance on all factors except the spatial memory factor (C3). In fact, the nonverbal cluster has a spatial memory performance equivalent to the high average cluster. Once again, as seen in Figure 18, the five cluster solution splits the average cluster into one cluster with slightly elevated general and spatial memory scores, an average forward recall score, and a somewhat diminished backwards recall score (C3) as well a second cluster with strong general memory and forward recall scores, a very strong backwards recall score, and an average spatial memory score (C2). From this it appeared that these two clusters differentiate primarily on their attention factor, with one cluster exhibiting strong attention scores and the other somewhat weaker scores. The six-cluster solution splits the low-average cluster so that there is now a cluster with low scores across all four factors (C1) and a cluster with somewhat average scores across the four factors (C2).

Cluster stability was evaluated with DFA. DFA correctly classified 89.4% of the four-cluster solution, 89.1% of the five-cluster solution, and 85.1% of the six-cluster solution - all comparable rates. The K-means iterative method then calculated clusters using means of the attributes as centroids and the cluster solutions derived this way were compared to Ward's method using Cohen's Kappa statistic. Kappa for the four-cluster solution was .67, for the five-cluster solution was .66, and for the six cluster solution was .12. From this it appears that the four and five-cluster solutions had good stability while the six-cluster solution had poor stability. See Tables 19, 20, and 21 for complete classification rates.

When the two-step clustering method was used to analyze the data, only two clusters emerged. As with the younger sample, the majority of participants fell into one cluster (97.1%) and only a few fell into the other cluster. Upon inspecting the minority cluster, it appears that children who scored uniformly low across all subtests made up this cluster. Given the lack of applicability to these findings, the two-step method was not further analyzed.

Beale's F statistic was then used with the four, five, and six cluster solution to see which cluster would be most appropriate to retain. As two comparisons were made, the Bonferroni correction set the alpha level at .025 (.05/2). Beale's statistic found that the five-cluster solution accounted for significantly more variance than the four-cluster solution, $F(14, 5208) = 3.11$, $p < .001$, but the six-cluster solution did not significantly account for more variance than the five-cluster solution, $F(14, 5194) =$

1.68, $p > .025$. Given that the six-cluster solution had very low stability, the five-cluster solution was therefore deemed the best solution for the 9-11 aged sample.

A comparison across gender on the five-cluster solution found a significant difference among clusters, $\chi^2(8) = 20.25$, $p < .01$. A breakdown of gender by cluster is available on Table 22.

As seen on the table, more boys fell into the nonverbal cluster (C3) and more girls fell in the high-average verbal and nonverbal yet average attention cluster (C4). Finally, a one-way ANOVA found no significant differences in age due to cluster membership, $F(5, 371) = .967$, $p > .05$. Prevalence rates were as follows: low average was 45%, elevated VMI/ACI was 28%, elevated VMI/NMI was 11%, elevated NMI was 9%, while high average was 7%.

Older Age Group (12-19 years)

As before, a four, five, and six cluster solution was each selected for the older sample. Tables 23, 24, and 25 have the means and standard deviations of standard scores in each cluster. These scores were plotted across the TOMAL VMI, NMI, and ACI as seen in Figures 19, 20, and 21. It is here that the advanced cluster with a superior ACI (SS ~ 140) emerged (C4), a pattern first identified in the entire standardization sample. A second cluster (C3) had high average across all three indexes. A third cluster captured average performance (C2) with a slight dip in the NMI and a fourth captured low average performance (C1) with a slight elevation in the NMI. Notably absent from this solution is a strong nonverbal cluster that appeared in the 9-11 aged group. With regard to the five-cluster solution, the relatively high functioning cluster split. The new cluster (C5) had a high average VMI and ACI but only an average NMI, suggesting this cluster to be more verbal in nature. Finally on the six-cluster solution the low average cluster split. The new cluster (C2) reflected a flat low average level of functioning while the previous low average cluster (C1) now reflects moderately low average functioning with a somewhat elevated (i.e. average) NMI.

Following this, these cluster solutions were plotted across the TOMAL standardization sample’s four factors. The four-cluster solution differs on the four factors by both level and pattern of performance. Specifically, the high performing cluster (C4) is high across all factors, and especially superior on the forward recall/attention factor. The low average cluster (C1) performed lower across all factors, while the average cluster (C2) is at the mean for three factors with a slight dip in the Spatial Memory factor. A fourth cluster (C3) has high average general memory performance and an elevated backward recall/working memory and spatial memory score, suggesting that this may be a cluster that represents participants with strong spatial and nonverbal abilities. In the five-cluster solution this particular cluster splits into two. One of the new clusters (C3) has the same pattern as before while the other cluster (C5) exhibits stronger scores on the Spatial Memory factor but below average scores on the Forward Recall/Attention factor. As seen in Figure 22, the six-cluster solution splits the low average cluster into two. These two clusters exhibit similar levels of performance on three of the factors and are only differentiated by the Spatial Memory factor, with one cluster (C1) performing near the mean and the other cluster (C2) performing four points below the mean.

DFA correctly classified 87.7% of the four-cluster solution, 88.5% of the five-cluster solution, and 88.5% of the six-cluster solution, indicating strong stability for all three clusters. Further analysis using the K-means iterative clustering method found that Cohen's kappa for the four-cluster solution was .81, for the five-cluster solution was .82, and for the six-cluster solution was .80. See Tables 26, 27, and 28 for cross-tabulation classification rates between the two clustering methods.

The two-step method only yielded a single cluster and was not further analyzed. Beale's F statistic was next used to see if the five or six-cluster solutions accounted for significantly more variance than the smaller solutions. The Bonferroni correction set the alpha level at .025 (.05/2). Results found that the five-cluster solution accounted for significantly more variance than the four-cluster solution, $F(14, 5292) = 6.57$, $p < .001$. In addition, the six-cluster solution accounted for significantly more variance than the five-cluster solution, $F(14, 5278) = 4.58$, $p < .001$. Based on these findings, the six-cluster solution was selected as the optimal solution for the 12-19 aged children.

A subsequent ANOVA among the six-cluster solution did find an overall significant effect for age, $F(5, 377) = 2.33$, $p < .05$, although post-hoc Scheffe tests found no significant differences among individual ages. Looking at the descriptives, it appears there is a trend in which the high performing cluster (C5) was about a year younger than the other clusters, with a mean age of 13.6, although this was not at a significant rate. Chi-square analysis found an effect for gender, $\chi^2(10) = 21.18$, $p < .05$. A breakdown of gender by cluster is available on Table 29. Prevalence rates were 19% for the low average cluster, 7% for the low cluster, 37% for the average cluster, 24% for the high average cluster, 9% for the advanced cluster, and 4% for the verbal cluster.

Final Analyses

The above analyses concluded that a five-cluster solution is optimal for the entire TOMAL standardization sample, a four-cluster solution for the 5-8 age group, a five-cluster solution for the 9-11 age group, and a six-cluster solution for the 12-19 age group. In order to provide further clinical information on these clusters' pattern of performances on the TOMAL, optimal cluster solutions were plotted across the TOMAL's other indexes: the SRI, FRI, LI, and ARI. Although these indexes share considerable overlap with the VMI, NMI, or ACI, they may capture additional constructs of memory and learning that may be further differentiated by cluster. First, the five-cluster solution of the TOMAL standardization sample was compared on these indexes as seen in Figure 23. As seen in the figure, three of the clusters remain steady on level of performance with a high average cluster (C4), an average cluster (C3), and a low average cluster (C1). Of the two remaining clusters, one has an average SRI but elevated LI, FRI, and ARI (C5), capturing individuals with average sequential recall but high average memory in other areas. As the SRI is very similar to content with the ACI, these are likely children who have average attentional skills. The other cluster (C2) had average performance on most indexes but a high performing SRI. This suggests that these individuals performed better than average on attentional tasks.

Next, the younger aged group (5-8 years) was plotted across these supplementary index scores using the four-cluster solution, as seen in Figure 24.

Consistent with the previous findings on the younger aged group, performance mostly differed by level and not by pattern. There was a high average cluster (C4), a low average cluster (C1), and two average clusters (C2 and C3). The two average clusters hovered near the mean, with one cluster (C3) slightly above and the other (C2) slightly below the mean. In addition, C2 exhibited a dip in the SRI, suggesting that these children may have poorer attentional capacities than average.

The mid-aged group (9-11 years) was best represented by a five-cluster solution and this solution was plotted across supplemental indexes in Figure 25. An advanced (C5), average (C2), and low average (C1) cluster emerged in this solution. Of note, the average cluster was actually slightly elevated for most of these indexes. Of the two remaining clusters, one (C3) represented children who performed average on the SRI but high average on the LI, FRI, and ARI to the point that their ARI scores matched the high performing cluster's ARI scores. This suggests these children have strong memory recall but average attention. The last cluster (C4) was unique in having average memory abilities on all indexes except the FRI. The FRI is composed of subtests that task children to freely recall stimuli without any prompting. Most of the FRI is composed of nonverbal subtests so it is likely that this is the high nonverbal cluster that was first identified in Figure 16.

Finally, the older group (12-19 years) was best represented by a six-cluster solution which was plotted across supplemental indexes in Figure 26. The low average (C2) and average (C3) clusters emerge once again. Of interest, the high performing cluster (C5) differs on pattern of performance, with a superior SRI score, strong FRI and LI, and somewhat less elevated ARI scores. This indicates that individuals in this cluster have very strong attention processing but are otherwise not too distinct from another cluster that had overall high average functioning (C4). A fifth cluster (C1) had lower than average performance although not as low as C2, and C1 also had slightly better FRI and LI scores than SRI and ARI scores. Finally, the last cluster (C6) was unique in having very distinct patterns across the four index scores, with low average SRI and LI, high average FRI, and high ARI scores. This suggests this cluster may have slower processing capabilities but overall good memorization techniques, with an additional advantage when presented with tasks that contain semantic content.

CHAPTER 5

DISCUSSION

The present study investigated the cluster profiles of children and adolescents as derived from a comprehensive battery of memory and learning. Age groups were found to split into 5-8 years, 9-11 years, and 12-19 years. An exploratory analysis on TOMAL raw scores as well as visual inspection of subtest patterns confirmed these specific age ranges. In order to account for the natural changes in memory that occur in children, these three age ranges were subsequently analyzed and compared. Cluster solutions were selected based on their stability, parsimony, and accounted variance. The four final solutions that were selected were compared across age and gender as well as supplemental TOMAL index scores. Overall findings fit with existing theory on memory development and appear to provide new insights on the nature of memory development in children and adolescents as measured by the TOMAL. A five-cluster solution was found to best describe the TOMAL standardization sample, which is not consistent with a previous study on TOMAL cluster patterns in children with TBI (Allen et al., in preparation). In their study, a four-cluster solution was a better fit for the matched controls that were compared to a TBI sample (which had a five-cluster solution). However, Allen and colleagues included the DRI in their analyses and had a significantly smaller sample size. Further, their sample may have been composed of younger children who fit a four-cluster solution in our current study.

Three, four, five, and six cluster solutions were considered when we examined cluster patterns of the entire standardization sample. One general finding that pervaded most solutions was that three of clusters that formed had a steady level of performance across all index and factor scores while other clusters differentiated general level of performance, as found in other studies. (Donders, 1996; Donders, 1999; Goldstein, Allen, & Seaton, 1999; Konold et al., 1999) Each of the steady level clusters represented children who performed at a certain level (advanced, high average, average, low average, low) suggesting that these children performed at a uniform level across most tasks. Given the shared variance between the TOMAL and intelligence batteries (Reynolds & Bigler, 1997), it is likely that the uniform levels of TOMAL performance reflect general intellectual abilities as well.

In this study, children in the high average/advanced clusters could be assumed to have stronger memory abilities and may also be children who are cognitively bright across several domains. On the other hand, children in the low average/low clusters have somewhat poorer (although not impaired) memory abilities and may also perform in the lower range in other domains. In addition, in some of the solutions an advanced cluster emerged with a superior ACI. This is consistent with the work by Allen and colleagues (in preparation) and suggests that exceptionally strong attentional capabilities may mediate performance on general memory in certain age ranges. In other words, these children were especially adept at focusing and sustaining attention and therefore performed well across all memory tasks.

The low average, average, and advanced clusters emerged on the three, four, five, and six cluster solutions indicating they consistently represented some of the children in the standardization sample. However, additional clusters emerged in the more complex cluster solutions that accounted for significantly more of the error variance. In the selected five-cluster solution, the two additional clusters represented alternative pattern of performance. One of these clusters captured children who had strong nonverbal performance (*M* = 114.4) as represented by the NMI, while still maintaining average VMI and ACI scores. Of interest, the strong NMI cluster also performed well on the LI and FRI, perhaps due to the shared nonverbal subtests in these supplemental index scores.

The other cluster that emerged had average VMI and NMI scores but a high average ACI score (*M* = 109.7). This cluster captured children with average memory skills but a distinct attention/recall ability that separated it from the other average cluster (this high average attention ability was also captured by the supplemental SRI). It appears that children in this cluster may have higher levels of working memory and can concentrate better on task performance, although this did not particularly impact their performance on other memory tasks. It is unclear why the attentional capacity did not appear to mediate memory performance on this cluster while it did on the high cluster, but it may be that when attentional capacity is superior it may increase general memory performance, but when it is simply elevated it does not. Conversely, it may be that attention may not directly influence memory performance but rather both memory and attention are both influenced by a third variable (i.e. *g*) which may elevate these two processes independently.

When plotted on TOMAL factor scores, the two clusters that differed by pattern had unique profiles. The NMI cluster had an elevated spatial memory factor (*M* = 12.7) that was close to the high functioning cluster. Of interest, this cluster also had a stronger working memory factor yet an average attention factor. This fits with previous findings which indicate that working memory is closely related to spatial memory, as both require the mental rotating and manipulation of information (Ramsey & Reynolds, 1995). On the other hand, the high average ACI cluster exhibited elevated working memory functioning (M = 12.1) while performing average in the general and spatial memory factors. In this case, it appears that working memory is not directly related to spatial memory.

One curious finding in subsequent comparisons of clusters was that the advanced cluster was slightly older than many of the other clusters, with the greatest difference at 1.2 years. Although standard scores were calculated to account for age, it appears that the high functioning cluster did represent more children who were older than other clusters. However, the age discrepancy is mild and perhaps best explained by the possibility that older children in the standardization sample performed exceptionally well compared to younger children. This may in part be because working memory continues to develop through adolescence (Gathercole, 1999). Therefore, even though standard scores correct for age, the advanced cluster captured older children who had strong working memory capacity simply by virtue of being older. This also provides an alternative explanation for the spike in the ACI relative to the VMI and NMI in the advanced cluster.

It should be noted that in the study of TOMAL cluster profiles in children with TBI (Allen et al., in preparation), the five clusters were different the standardization sample. In their study, the highest functioning cluster actually approached the population average (M = 100) on all the indexes. A second cluster had particularly poor performance across all indexes, scoring more than two standard deviations below the mean. A third cluster had average performance on most indexes but a poor ACI, while two other clusters had below average performance on most indexes although one had stronger VMI scores and another stronger NMI scores. These contrasts are important to note, as TBI is heterogeneous in mechanism and outcome of injury (Mottram & Donders, 1996). In Allen and colleagues' study, TBI appears to impair attention and concentration on some children while selectively damaging verbal or nonverbal memory domains in other children, perhaps due to unilateral lesions in the hippocampus. In the present study, findings indicate that children and adolescents perform high, average, low average, average with high NMI, and average with high average ACI on the TOMAL. This natural memory variation may in part account for memory performance and cluster membership in children who sustain TBI. For example, it is possible that children who were previously high functioning fall into the average cluster post-injury. Allen and colleagues' study was almost limited by the fact that they could not separately cluster TBI age groups due to their limited sample size (n = 150). A subsequent study could estimate premorbid functioning in TBI children and subsequently predict whether cluster membership prior to injury may predict cluster membership post-injury.

After establishing the optimal cluster solution of the entire standardization sample, the sample was then broken down into age groups. The review by Gathercole (1998) concludes that several domains of memory, including short-term, visuospatial, episodic, and executive functioning make rapid gains from infancy through about age 7. Around age 8, however, memory appears to reach adult-levels in organization and strategies and improvement generally steadies over age and is fully developed around age 12 (Schneider & Pressley, 1997). Gathercole also notes that different domains seem to slow at different ages; for example, short-term memory appears to steady after age 7, while visuospatial appears to level off around age 12 and executive functioning/working memory continues to develop throughout adolescence. Another study by Gathercole (1999) looked at performance across several domains of memory in children and confirmed that performance seems a sharp increase from ages 5-8 after which it flattens. The one exception to this is working memory tasks which rise steadily through young adulthood.

As previously discussed, it was apparent that the sample is best explained by three age groups (5-8, 9-11, 12-19 years). We came to this conclusion due to a number of reasons. First, visual inspection of TOMAL raw subtest performance plotted across age indicated that the two age points where memory performance exhibited the biggest jumps were at ages 9 and 12 respectively. Next, an exploratory two-step cluster analysis using raw scores as attributes yielded three clusters with mean ages of 7.1, 11.1, and 12.9 (*SD* = 1.8, 2.7, 2.9. A second cluster analysis using Ward's method and specifying three clusters found mean ages to be 8.2, 11.5, and 13.3 (*SD* =2.7, 3.0, 2.9). These mean ages roughly fit within the three proposed age groups that appeared on the plots. Finally, a review of the literature indicated that short-term phonological memory development appears steady off around age 8 while short-term visuospatial memory, as well as general memory organization and strategies steady off around age 12 (Schneider & Pressley, 1997), with working memory continuing to improve through adulthood.

The 5-8 years group represents our youngest participants and reflects a period of rapid memory development and growth. TOMAL standard scores removed the effects of age from cluster membership and allowed us to investigate memory patterns based on average performance and not on natural changes with age. As before, a four, five, and six cluster solution was proposed. However, much of the analyses was halted because very few (i.e. six) participants fell into the additional clusters, making these clusters difficult to interpret or generalize. A cursory examination of the additional clusters suggested they captured kids who were exceptionally gifted and scored far and beyond the expected range of performance. Due to the extremely small membership of the clusters, a three-cluster solution was also examined. However, the four-cluster solution accounted for more variance and was selected as the optimal solution.

Unlike the standardization sample, the 5-8 years group was mostly differentiated by level of performance with no clusters particularly deviating on pattern of performance. Two two clusters represented children who scored high average and low average uniformly across TOMAL index scores. Less pronounced in the highest performing cluster was a unique elevation in the ACI; it appears in younger groups that attention is not selectively superior in the already high average/advanced cluster. Unique to the younger age group, there was no strictly "average" cluster per se but rather two clusters that hovered above and below the population mean respectively. One cluster had slight elevations in the VMI and ACI with a slight dip towards 100 in the NMI, while the other cluster had slight deviations in the VMI and ACI with a slight elevation towards 100 in the NMI. Although the NMI appeared to slightly differ on pattern of performance for the two average clusters, essentially there were no marked deviations in index score performance; as with the high average and low average clusters, these average clusters were more defined by level rather than pattern of performance.

Inspection of the four-cluster solution across supplemental index scores suggests that the supplemental scores had relatively flat pattern of performance as well. Factor scores exhibited similar pattern of performance, although it is interesting to note that in all four clusters, the working memory factor exhibited a mild elevation relative to the other factors. It is important to note that young children as a whole generally scored better on backward recall tests compared to forward recall tests. This pattern may be partially explained findings which suggest that forward and backward span tasks tap into different aspects of memory. It has been argued that forward span tasks tap into short-term auditory memory, verbal sequencing, and simple verbal expression (Hale, Hoeppner, & Fiorello, 2002) while backwards span tasks load more heavily onto working memory (Ramsay & Reynolds, 1995). Gathercole (1998) notes the different developmental rates of short-term memory and working memory and this may account for the differential pattern of performance between TOMAL forward and backward span tasks.

When the clusters were compared across age and gender, no significant differences were found. Therefore, the present study's findings indicate that memory in younger children is mostly defined by ability and not by specific strengths or weaknesses within subcomponents. Verbal and nonverbal domains develop at a relatively constant rate (whether at a high, middle, or low functioning rate). It is worth nothing that attention also develops at a constant rate, although working memory appears to develop slightly faster compared to straightforward attention span tasks.

In contrast to the 5-8 years group, the 9-11 years group was classified best with a 5-cluster solution. In addition, three of the clusters differed on pattern of performance while the other two remained differentiated by level of performance. Perhaps most striking about this cluster solution is that no single cluster truly represented an average level of performance. Although there was a high average and low average cluster which performed at a steady level across the VMI, NMI, and ACI, the other three clusters had greater variation across these three index scores. As found in the entire standardization sample, one cluster that emerged had a strong NMI (*M* = 117.3) compared to an average VMI and ACI. It is apparent that the children in this cluster have a particular strength in nonverbal memory tasks while performing average on tests of verbal memory and attention. Another cluster captured children who performed above average on the VMI (*M* = 110.7) and NMI (*M* = 111.4) but a little below average on the ACI (*M* = 95.0). These children apparently have strong general memory capabilities but only average attention capacity. A third cluster had high average VMI (*M* = 109.2) and ACI (*M* = 112.4) scores but only average NMI scores (*M* = 98.4). This cluster captured children who do particularly well on verbal tasks. Of note, the content of the attention subtests are largely verbal by nature, so it is possible that this is a cluster with stronger verbal performance which also is reflected by the ACI.

In some ways the 9-11 years cluster pattern resembles the standardization sample's cluster pattern the most, although some differences are also apparent. For one, the advanced cluster does not exhibit a particularly elevated ACI which appeared on the standardization sample; as suggested earlier, this could be due to the fact that working memory continues to develop throughout adolescence and young adulthood. In other words, older children (as captured by the standardization sample) more frequently fit into the advanced cluster, which had a spike in the ACI, because of their natural working memory development. In addition, the entire standardization sample had an average cluster, while this cluster was replaced by the high average VMI/NMI, average ACI cluster in the 9-11 years group. The period of memory development between 9 and 11 is a critical time for children, as general memory development steadies and organizational strategies begin to solidify (Schneider & Pressley, 1997). It is possible that the greater variation in performance across clusters represents children who are developing their strategies at different rates.

The 9-11 years group demonstrated similar variation on the TOMAL factor scores. Particularly striking is the nonverbal cluster, which had typical performance on the attention and working factors and an elevation in the spatial memory factor (M = 14.9) which approached that of the high average cluster. This indicates that the nonverbal cluster is particularly strong in spatial memory tasks and near the level of the high average cluster that performed uniformly well across all factors and indexes. The strong ACI cluster appears to have a particular strength in working memory (M = 12.6) and typical performance on general and spatial memory tasks. In contrast, the strong VMI/NMI cluster had typical performance on the attention factor and below average performance (M = 8.9) on the working memory factor. This counteracts the findings on the entire standardization sample which had evidenced that working memory may mediate spatial memory tasks, as some children appear strong in working memory and average in spatial memory, while others are strong in spatial memory and weak in working memory. However, as this cluster only focuses on the 9-11 years group, it may be that when all ages are considered, working memory and spatial memory do covary to a degree that is not as expressed in this particular age range.

Supplemental index scores can be difficult to interpret, as they share considerable overlap with the VMI, NMI, and ACI. As expected, the high average cluster performed well across all supplemental indexes while the low average cluster performed poorly. The high VMI/ACI cluster had elevations across the SRI and average performance on the LI and FRI, which are two indexes that share nonverbal subtest. Of note, this cluster also performed near the average range on the ARI, which is an index composed of subtests of semantic verbal recall. It appears this cluster performed better on verbal subtests unrelated to semantic recall, such as word selective reminding and object recall. The strong NMI cluster performed very well on the FRI index, although this is likely explained by the FRI primarily being composed of nonverbal subtests. Finally, the strong VMI/NMI cluster performed well on all supplemental indexes except the SRI, which shares subtests with the ACI.

The 9-11 years clusters did not differ across age. However, the nonverbal cluster did have significantly more males than females. Gender differences in spatial abilities have been documented (Halpern, 2000; Robinson et al., 1996; Voyer, Voyer, & Bryden, 1995) and prior research has indicated that boys tend to do better on TOMAL spatial tasks (Lowe, Mayfield, & Reynolds, 2003). Therefore, it is unsurprising that there would be gender differences in cluster membership, although it is interesting that the differences appear very pronounced in the 9-11 years group.

The 12-19 years group fit a 6-cluster solution, providing further evidence that as children develop, their cognitive processes grow more complex and diversified. This supports research that indicates that memory abilities become more differentiated with age as developmental differentiation of neural circuits and abilities emerge (Thomasen et al., 2009). It is within this age range that the advanced cluster displays a marked elevation in the ACI which is later seen in the entire standardization sample. It appears that the superior ACI may be because working memory develops throughout adolescence, and working memory is represented by the ACI. Even though scaled scores are age-corrected, it may be that older children (i.e. 18 and 19 year olds) fell more frequently in the advanced cluster based on their advanced working memory. In contrast, the memory components represented by the VMI and NMI do not increase as markedly in older children. Therefore, in the advanced cluster, some of the adolescents who had superior ACI scores had less advanced VMI and NMI, accounting for the lower (although still advanced) scores in these two indexes.

Two more clusters had even level of performance; one solution emerged as low as all index scores fell 20 points below the mean. Unique to the older children, a high-average cluster emerged distinct from the advanced cluster with scores about 10 points above the mean across the three indexes. The remaining three clusters varied on pattern as well as level of performance. Once cluster had strong VMI scores but average NMI and slightly elevated ACI scores. As posited with the 9-11 age group, this cluster’s ACI may be influenced by a strong verbal ability, given the verbal nature of ACI subtests. This cluster then is best categorized as a verbal cluster. The fifth cluster hovered around the population mean with a slight dip in the NMI. Although the dip in the NMI is noticeable, this cluster may best represent children who performed average on the TOMAL. The sixth cluster captured children who had mediocre VMI scores, poor ACI scores, and near average NMI scores. This low-average cluster did not function as poorly as the low cluster, but did perform below the mean on all indexes.

When the clusters were plotted across the factor scores, more information is provided. Most striking, the strong verbal cluster actually scored extremely well on the spatial memory factor (M = 13.4). This cluster appears to be the equivalent of the strong spatial memory cluster on the 9-11 years range, although in the 12-19 years range, spatial memory separates from nonverbal memory in this cluster. In other words, these children did very well in spatial memory and verbal memory, but performed only at average in other nonverbal memory tasks.

As expected, the advanced cluster did very well across the four factors, with strong scores on the attention and working memory factors. Similarly, the low cluster had poor performance on all factors with a greater dip in the spatial memory factor. Of interest, the low-average functioning cluster had similar poor attention and working memory scores as the low cluster, but actually had an increase in spatial memory. It seems that spatial memory alone differentiated these two clusters and identified children with overall weak memory and attention capabilities yet strong or weak spatial memory ability. The average cluster performed average on the factors with a slight dip in spatial memory and the high-average functioning appears to have particular strengths in working and spatial memory. In looking at the 6-cluster solution factor scores, it seems that spatial memory ability differentiates the clusters the most, followed by working memory and attention abilities.

Finally, the clusters were plotted across supplemental index scores. Once again, the verbal/spatial memory cluster showed the most unique pattern profile. For one, the SRI was actually quite low in this cluster. The SRI is composed of forward sequential tasks, including a nonverbal sequential subtest (VSM). It seems that children in this cluster performed poorly in forward sequential tasks. However, this cluster outperformed even the advanced cluster on the ARI, an index of semantic verbal recall (MFS and PR). This may account for the strong VMI score as well. It appears that this one cluster has particularly selective strengths and weaknesses that are not best explained by the VMI, NMI, and ACI alone. Rather, this cluster has overall strong spatial memory, free recall, and semantic recall abilities and weak forward sequential capacity, while performing somewhat average in learning.

The other clusters had more predictable pattern of performance on the supplemental indexes. The advanced cluster remains elevated on most of the indexes, with SRI scores (matching the ACI score). The high-average, average, and low clusters performed at flat levels across all indexes. Finally, the low-average cluster appeared to have stronger FRI and LI scores than ARI and SRI, although this is likely accounted for by the shared nonverbal subtests in the former two indexes. Despite this, the low-average cluster did perform below the mean on all indexes, regardless of specific variations.

These findings unequivocally highlight the diverse memory components captured by the TOMAL. Consistent across all age groups, advanced, average, and low/low average functioning clusters emerged in our sample. In addition, some clusters had selective differences on the verbal, nonverbal, and attention/concentration indexes. It appears that the NMI especially separates from the VMI and ACI. For example, in the entire standardization sample and in the 9-11 years group, a very strong NMI cluster emerged. The TOMAL has been praised for having strong construct validity in assessing nonverbal domains (Miller, Bigler, & Adams, 2001). Judging by the derived cluster solutions, this appears to be the case; the NMI frequently diverged from the VMI and ACI.

The ACI also showed unique pattern of performance on specific clusters. For example, the advanced cluster in older children and in the standardization sample was shown to have an even higher ACI, although as suggested, this may be due to the natural development of working memory in adolescents approaching young adulthood. The ACI also was strong in a specific cluster in the standardization sample and weak in a cluster in the 9-11 years group. Prior factor studies on the TOMAL have consistently extracted an attention factor (Alexander & Mayfield, 2005; Allen et al., in preparation, Reynolds & Bigler, 1995). The ACI has also been showed to be sensitive in children with attention deficit/hyperactivity disorder (Thaler et al., 2009). The findings in the present paper further add strength to the construct validity of the ACI, as an index that repeatedly and reliably separates from other indexes and measures unique processes in children and adolescents.

Attention is further examined via the factor structure where the four span tests are further split into forward and backward factors. Clusters in the various age ranges had differential performance on these factors. In many cases, the backward span (e.g.

working memory) factor appeared to share some variance with the spatial memory factor while the forward span (e.g. attention) factor appears to vary less with the spatial memory factor and more with the general memory factor. The spatial memory factor itself appears to capture a construct of its own not best represented by the NMI: this is evidenced in the older age range which had a cluster with only average NMI yet superior spatial memory scores. It is worthwhile for clinicians to consider the TOMAL factors in assessing their patients, with particular emphasis on the spatial memory factor followed by the working memory and attention factors.

Overall, the supplemental index scores provided little information due to their shared overlap with the core index scores. For example, the FRI purportedly measures free recall but is mostly made up of subtests with a nonverbal modality; in this way, clusters with high NMI typically had higher FRI scores as well. In a similar vein, the SRI and ACI share most of their subtests as well. Perhaps the most illuminating of the supplemental index scores is the ARI, as it is composed of only two verbal subtests that have unique composition: the memorization and recall of cued verbal information (such as story recall or associated word recall). A cluster in the 12-19 group appeared to have unique performance on the ARI, providing some information about this cluster's participants and their semantic recall abilities.

In tying these findings with cognitive components, the cluster solutions appear to consistently show stable verbal, nonverbal, forward recall, backward recall, and spatial memory components with the possible addition of semantic verbal recall. These components appear to be stable because 1) they consistently emerge as separable components that vary within cluster performance and 2) they appear to match the existing literature on known details of these components (e.g. working memory's relationship with spatial memory, nonverbal abilities in boys vs. girls). Clinicians would do well to make these components areas of specific focus in evaluating children and adolescents with the TOMAL. Further, clinicians can use the provided cluster profiles to posit their patients' cluster membership and general abilities across several memory domains. Other memory components measured by the TOMAL such as learning and free recall are more difficult to interpret due to these indexes significant overlap with core indexes.

Prevalence of clusters generally suggests that cluster represent an adequate proportion of their respective age groups. Generally, children fell within average or low average groups more frequently than high average groups. Average scores may be calculated then, not by equivalent low average/high average variances which pull around the middle but rather by the additional pattern of performance clusters which contribute to more positive variance compared to negative variance. The six-cluster solution for older children did have a verbal cluster that only accounted for a small part of the sample (e.g. 4%) but in most cases, between 10-40% of each cluster solution's participants were captured by a cluster, thereby establishing that these cluster profiles occur frequently enough to warrant notice by assessors. In addition, cluster patterns that fall out of the established clusters may be considered extraordinary and perhaps require clinical attention.

In analyzing the clusters, Ward's minimum variance method was universally the preferred clustering method. In all cases, the two-step method simply did not account for enough variance in participants and grouped the sample into either one or two-cluster solutions, with the two-cluster solutions often heavily imbalanced. This may be in part because the two-step cluster has different criteria in selecting sub-clusters than in calculating final cluster solutions (Zhang, Ramakrishnon, & Livny, 1996). In selecting sub-clusters, the two-step method absorbed many of the cases into only a few sub-clusters, which then were clustered into a single solution in the second step. In any case, the two-step method did prove useful when clustering raw scores, as it found a three-cluster solution that matched the three age groups we had originally hypothesized. It may be that when differences in subtest scores are less pronounced, the two-step method merges them more than Ward's method. When age is unaccounted for (as in raw scores), subtest scores expressed a great deal more variation which ultimately was parsed into three clusters. When age is corrected (as in scaled scores), subtest variation is not enough for the two-step method to reliably differentiate cluster membership.

Limitations

Some limitations need to be addressed. First, the TOMAL has a delayed recall index (DRI) that was not accounted for in the current study. The DRI was excluded because of its limited range in subtest scores, as suggested by Reynolds and Bigler (1996). However, the DRI should be further investigated in other studies, as memory consolidation has been identified as a cognitive component (Golimbet et al., 2006). Therefore, the relationship between the DRI and other TOMAL index and factor scores is still unknown.

Another limitation is that although reliable and stable clusters were extracted from the study, these clusters' applicability to other variables is yet unknown. As all the analysis took part within the standardization sample, only age and gender were compared across the clusters. It would be very helpful to run additional studies comparing these clusters across other cognitive, emotional, and behavioral measures in order to better arrive at meaningful conclusions about these clusters. As it stands, the results seem to suggest that the clusters are helpful in identifying normal patterns of memory variation. Donders (1996) makes the point that standardization sample cluster analyses are useful because the results can identify normally co-occurring profiles that can then be used for comparison to profiles obtained in clinical populations, in order to determine if the clinical profiles are unusual.

In summary, the standardization sample had 5 clusters and this sample was best divided into the hypothesized 5-8 years, 9-11 years, and 12-19 years ranges. Further, clusters in the different age groups increased in number and complexity, with younger children having relatively flat level of performance while the oldest children had subtle variations that were captured by factor and supplemental indexes along with the core indexes. Future studies should compare additional clinical samples with this standardization sample as well as further investigate the implications of cluster membership across other assessment variables. By further researching these clusters and comparing them to clusters that emerge in clinical samples, we will further understand the delicate and intricate variations of memory exhibited by developing children and adolescents.

REFERENCES

Abel, T., Martin, K. C., Bartsch, D., & Kandel, E. R. (1998). Memory suppressor genes: inhibitory constraints on the storage of long-term memory. *Science, 279*, 338-341.

Aldenderfer, M., & Blashfield, R. (1984). *Cluster Analysis.* Sage Publications, Newbury Park.

Alexander, A. I., & Mayfield, J. (2005). Latent factor structure of the Test of Memory and Learning in a pediatric traumatic brain injured sample: Support for a general memory construct. *Archives of Clinical Neuropsychology, 20,* 587-598.

Allen, D. N., Leany, B. D., Thaler, N. S., Cross, C, & Mayfield, J. (in preparation). Parsing neurocognitive heterogeneity in pediatric traumatic brain injury.

Anderberg, M R. (1973). *Cluster Analysis for Applications.* New York: Academic Press. Ando, J., Ono, Y., & Wright, M. J. (2001). Genetic structure of spatial and verbal working memory. *Behavioral Genetics, 31,* 615-624.

Bacher, J., Wenzig, K., & Vogler, M. (2004). SPSS two-step cluster – a first evaluation. In *RC33 sixth international conference on social science methodology: Recent developments and applications in social research methodology, Amsterdam, the Netherlands.*

Baddeley, A.D. & Hitch, G.J. (1974). Working memory. In G.A. Bower (ed.), *Recent Advances in Learning and Motivation*, Vol. 8 (pp. 47–89). New York: Academic Press.

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

Bahrick, H. (2000). Long-term maintenance of knowledge. *The Oxford handbook of memory* (pp. 347-362). New York, NY US: Oxford University Press.

Bauer, P. J. (1996). What do infants recall of their lives? Memory for specific events by one to two-year-olds. *American Psychologist, 51*, 29-41.

Bawden, H. N., Knights, R. M., & Winogron, H. W. (1985). Speeded performance following head injury in children. *Journal of Clinical and Experimental Neuropsychology, 7,* 39-54.

Beale, E. M. (1969). Euclidean cluster analysis, in *Bulletin of the International Statistical Institute: Proceedings of the 37th Session (London),* Book 2, pp. 92-94, Voorburg, Netherlands:ISI.

Binder, D. K., & Scharfman, H. F. (2004). Brain-derived neurotrophic factor. *Growth Factors, 22,* 123-131.

Binet, A. (1907). *A practical guide for the admission of retarded children into special classes.* Paris: Colin.

Brzustowicz, L. M., & Bassett, A. S. (2008). Phenotype matters: The case for careful characterization of relevant traits. *American Journal of Psychiatry, 165*(9), 1096-1098.

Clatworthy, J., Buick, D., Hankins, M., Weinman, J., & Home, R. (2005). The use and reporting of cluster analysis in health psychology: A review. *British Journal of Health Psychology, 10,* 329-358.

Corder, E. H., Saunders, A. M., Strittmatter, W. J., Schmechel, D. E., Gaskell, P. C., Small, G. W., et al. (1993). Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer’s disease in late onset families. *Science, 261,* 921-923.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297-334.

Davis, M., Hitchcock, J. M., & Rosen, J. B. (1987). Anxiety and the amygdala: Pharmacological and anatomical analysis of the fear-potentiated startle paradigm. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 21, pp. 263–305). New York: Academic Press.

Delis, D., Kramer, J., Kaplan, E., & Ober, B. (1994). *California Verbal Learning Test – Children's Version.* San Antonio, TX: Psychological Corporation.

Dempster, E., Toulopoulou, T., McDonald, C., Bramon, E., Walshe, M., Filbey, F., et al. (2005). Association between BDNF val66 met genotype and episodic memory. *American Journal of Medical Genetics B Neuropsychiatric Genetics, 134*, 73-75.

Donders, J. (1996). Cluster subtypes in the WISC-III standardization sample: Analysis of factor index scores. *Psychological Assessment*, *8*(3), 312-318.

Donders, J. (1999). Cluster subtypes in the standardization sample of the California Verbal Learning Test – Children's Version. *Developmental Neuropsychology, 16*(2), 163-175.

Donders, J. (2008). Subtypes of learning and memory on the California Verbal Learning Test-Second Edition (CVLT-II) in the standardization sample. *Journal of Clinical and Experimental Neuropsychology*, *30*(7), 741-748.

Donders, J., & Warschausky, S. (1997). WISC-III factor index score patterns after Traumatic Head Injury in Children. *Child Neuropsychology, 3*(1), 71-78.

Donders, J., Zhu, J., & Tulsky, D. (2001). Factor index score patterns in the WAIS-III standardization sample. *Assessment, 8,* 193-203.

Doyle, A. E., Willcutt, E. G., Seidman, L. J., Biederman, J., Chouinard, V., Silva, J., & Faraone, S. V. (2005). Attention-deficit/hyperactivity disorder endophenotypes. *Biological Psychiatry, 57*(11), 1324-1335.

Ferris, L. M., & Kamphaus, R. W. (1995). Review of the Test of Memory and Learning. *Archives of Clinical Neuropsychology, 10*(6).

Finkbeiner, S. (2000). Calcium regulation of the brain-derived neurotrophic factor gene. *Cellular & Molecular Life Sciences, 57,* 394-401.

Frantom L. V., Allen, D. N., & Cross, C. (2008). Neurocognitive endophenotypes for bipolar disorder. *Bipolar Disorders, 10,* 387-399.

Fuller, K. M. (2001). The relationship of attention, memory, and academic achievement in children with ADHD. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 61*(10-B), 5562.

Gathercole, S. E. (1998). The development of memory. *Journal of Child Psychology and Psychiatry, 39*(1), 3-27.

Gathercole, S. E. (1999). Cognitive approaches to the development of short-term memory. *Trend in Cognitive Science, 3*(11), 410-419.

Gathercole, S. E., & Baddeley, A. D. (1989). Evaluation of the role of phonological STM in the development of vocabulary in children: A longitudinal study. *Journal of Memory and Language, 28,* 200-213.

Gathercole, S. E., & Hithc, G. J. (1993). Developmental changes in short-term memory: A revised working memory perspective. In A. Collins, S. E. Gathercole, M. A. Conway, & P. E. Morris (Eds.), *Theories of memory* (pp.

189-210). Hove, U.K.: Erlbaum.

Gazzaniga, M.S. (1995) *The Cognitive Neurosciences*. Cambridge, MIT Press.

Glutting, J. J., McDermott, P. A., Prifitera, A., & McGrath, E. A. (1994). Core profile types for the WISC-III and WIAT: Their development and application in identifying multivariate IQ-achievement discrepancies. *School Psychology Review, 23,* 619-639.

Goldstein, G., Allen, D. N., & Seaton, B. E. (1998). A comparison of clustering solutions for cognitive heterogeneity in schizophrenia. *Journal of the International Neuropsychological Society, 4,* 353-362.

Golimbet, V., Alfimova, M., Kaleda, V., Lavrushina, O., Korovaitseva, G., Lezheiko, T., et al. (2006). Verbal memory deficit in schizophrenia as a possible endophenotype of the disease. *Schizoaffective Disorders: New research* (pp. 165-186). Hauppauge, NY US: Nova Science Publishers.

Gottesman, I., & Gould, T. (2003). The endophenotype concept in psychiatry: Etymology and strategic intentions. *American Journal of Psychiatry, 160,* 636-645.

Hale, J. B., Hoeppner, J. B., & Fiorello, C. A. (2002). Analyzing Digit Span components to assessment of attention processes. *Journal of Psychoeducational Assessment, 20* (2), 128-143.

Halpern, D. F. (2000). *Sex differences and cognitive abilities.* Mahwah, NJ: Erlbaum.

Hartman, D. E. (2007). Test review: PSST! Wanna buy a good new memory test – cheap?: The Test of Memory and Learning-2. *Applied Neuropsychology, 14*(4), 307-309.

Hebb, D. O. (1949). *The Organization of Behavior: A Neuropsychological Theory*. New York: Wiley.

Heijden, P., & Donders, J. (2003). WAIS-III factor index score patterns after traumatic brain injury. *Assessment, 10*(2), 115-122.

Ho, B. C., Milev, P., O'Learny, D. S., Librant, A., Andreasen, N. C., & Wassink, T. H. (2006). Cognitive and magnetic resonance imaging brain morphometric correlates of brain-derived neurotrphic factor Val66Met gene polymorphism in patients with schizophrenia and healthy volunteers. *Archives of General Psychiatry, 63,* 731-740.

Howes, N. L., Bigler, E. D., Lawson, J. S., & Burlingame, G. M. (1999). Reading disability subtypes and Test of Memory and Learning. *Archives of Clinical Neuropsychology, 14*(3), 317-339.

Kaufman, A. S., & Lichtenberger, E. O. (1999). *Essentials of WAIS-III assessment.* New York: John Wiley.

Kaufman, L., & Rousseeuw, P. J. (1990). *Finding groups in data: An introduction to cluster analysis.* Wiley-Interscience.

Kliegel, M., Mackinlay, R., & Jager, T. (2008). Complex prospective memory: Development across the lifespan and the role of task interruption. *Developmental Psychology, 44*(2), 612-617.

Kremen, W. S., Jacobsen, K. C., Xian, H., Eisen, S. A., Eaves, L. J., Tsuang, M. T., et al. (2007) Genetics of verbal working memory processes; A twin study of middle-aged men. *Neuropsychology, 21,* 569-580.

Lange, R. T. (2007). WAIS-III index scores profiles in the Canadian standardization sample. *Journal of Clinical and Experimental Neuropsychology, 29*(1), 47-58.

Lajiness-O'Neill, R., Beaulieu, I., Titus, J. B., Asamoah, A., Bigler, E. D., Bawle, E. V., & Pollack, R. (2005). Memory and learning in children with 22q11.2 deletion syndrome: evidence for ventral and dorsal stream disruption? *Child Neuropsychology, 11,* 55-71.

Lowe, P. A., Mayfield, J. W., & Reynolds, C. R. (2003). Gender differences in memory test performance among children and adolescents. *Archives of Clinical Neuropsychology, 18,* 865-878.

Lowther, J., & Mayfield, J. (2004). Memory functioning in children with traumatic brain injuries: A TOMAL validity study. *Archives of Clinical Neuropsychology, 19*(1), 89-104.

Lynch, G., & Baudry, M. (1984). The biochemistry of memory: a new and specific hypothesis. *Science, 224*(4653), 1057-1063.

Lynch, G., Rex, C. S., & Gall, C. M. (2006). Synaptic plasticity in early aging. *Ageing Research Review, 5,* 255-280.

MacQueen, J. B. (1967). Some methods for classification and analysis of multivariate observations. *Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability. University of California Press,* 281-297.

McCabe, J., & Hartman, M. (2008). Working memory for item and temporal information in younger and older adults. *Aging, Neuropsychology, and Cognition, 15,* 574-600.

McIntyre, R. M., & Blashfield, R. K. (1980). A nearest-centroid technique for evaluating the minimum variance clustering procedure. *Multivariate Behavior Research, 15,* 225-238.

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review, 101*(2), 343-352.

Miller, M. J., Bigler, E., & Adams, W. V. (2001). Clinical neuropsychological assessment of child and adolescent memory with the WRAML, TOMAL, and CVLT--C. In A. S. Kaufman & N. L. Kaufman (Ed.), *Specific Learning Disabilities and Difficulties in Children and Adolescents : Psychological Assessment and Evaluation* (pp. 387-429), Cambridge: University Press.

Milner, B., Squire, L. R., & Kandel, E. R. (1998). Cognitive neuroscience and the study of memory. *Neuron, 20,* 445-468.

Mitchell, W. G., Chavez, J. M., Baker, S. A., Guzman, B. L., & Azen, S. P. (1990). Reaction time, impulsivity, and attention in hyperactive children and controls: A video game technique. *Journal of Child Neurology, 5,* 195-204.

Mondadori, C. R., deq Uervain, D. J., Buchmann, A., Mustovic, H., Wollmer, M. A., Schmidt, C. F., et al. (2007). Better memory and neural efficiency in young apolipoprotein E epsilon4 carriers. *Cerebral Cortex, 17,* 1934-1947.

Morrison, J. R. (2006). Discriminant function analysis using the TOMAL: Comparison between TBI, ADHD, and control groups of children and adolescents. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 67,* 2272.

Mottram, L., & Donders, J. (2005). Construct validity of the California Verbal Learning Test – Children's Version (CVLT-C) after pediatric traumatic brain injury. *Psychological Assessment, 17,* 212-217.

Mottram, L., & Donders, J. (2006). Cluster subtypes on the California verbal learning test-children's version after pediatric traumatic brain injury. *Developmental Neuropsychology, 30*(3), 865-883.

Murer, M. G., Yan, Q., & Raisman-Vozari, R. (2001). Brain-derived neurotrophic factor in the control human brain, and in Alzheimer's disease and Parkinson's disease. *Progress in Neurobiology, 63,* 71-124.

Murji, S., Rourke, S. B., & Donders, J. (2003). Theoretically derived CVLT subtypes in HIV-1 infection: Internal and external validation. *Journal of the International Neuropsychological Society, 9,* 1-6.

Okura, T. T. (2001). The validity of the construct of verbal and nonverbal memory in various measures. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 62*(1-B), 559.

Old, S. R., & Naveh-Benjamin, M. (2008). Differential effects of age on item and associative measures of memory: A meta-analysis. *Psychology and Aging, 23*(1), 104-118.

Packard, C. J., Westendorp, R. G., Stott, D. J., Caslake, M. J., Murray, H. M., Shepherd, J., et al. (2007). Propsective study of pravastatin in the elderly at risk group. Association between apolipoprotein E4 and cognitive decline in elderly adults. *Journal of the American Geriatric Society, 55,* 1777-1785.

Payton, A. (2006). Investigating cognitive genetics and its implications for the treatment of cognitive deficit. *Genes, Brain, & Behavior, 5,* 44-53.

Prigatano, G. P. (1990). Recovery and cognitive retraining after cognitive brain injury. In E. Bigler (Ed.), *Traumatic brain injury* (pp. 273-295), Austin, TX: PRO-ED.

Rao, N. P., Reddy, Y. C. J., Kumar, K. J., Kandavel, T., & Chandrashekar, C. R. (2008). Are neuropsychological deficits trait markers in OCD? *Progress in Neuro-psychopharmacology & Biological Psychiatry, 32,* 1574-1579.

Ramsey, M. C., & Reynolds, C. R. (1995). Separate Digits tests: A brief history, a literature review, and a reexamination of the factor structure of the test of memory and learning (TOMAL). *Neuropsychology Review, 5*(3), 151-171.

Raz, N., Rodrigue, K. M., Kennedy, K. M., & Land, S. (2009). Genetic and vascular modifiers of age-sensitive cognitive skills: effects of COMT, BDNF, ApoE, and Hypertension. *Neuropsychology, 23*(1), 105-116.

Reeves, D., & Wedding, D. (1994). *The clinical assessment of memory.* Berlin: Spring-Verlag.

Reynolds, C. R., & Bigler, E. D. (1994). *Test of Memory and Learning: Examiner's Manual*. Austin, TX: PRO-ED.

Reynolds, C. R., & Bigler, E. D. (1994b). *Manual for the Test of Memory and Learning.* Austin, TX: PRO-ED.

Reynolds, C. R., & Bigler, E. D. (1996). Factor structure, factor indexes, and other useful statistics for interpretation of the Test of Memory and Learning (TOMAL). *Archives of Clinical Neuropsychology, 11*(1), 29-43.

Reynolds, C. R., & Bigler, E. D. (1997). Clinical neuropsychological assessment of child and adolescent memory with the Test of Memory and Learning. In C. Reynolds & E. Fletcher-Janzen (Ed.), *Handbook of Clinical Child Neuropsychology* (2nd ed, pp. 296-319). New York: Plenum.

Reynolds, C. R., & Voress, J. K. (2007). *Test of Memory and Learning: Second Edition.* Austin, TX: PRO-ED.

Robinson, N. M., Abbott, R. D., Berninger, V. W., & Busse, J. (1996). The structure of abilities in math-precocious young children: Gender similarities and differences. *Journal of Educational Psychology, 88*(2), 341-352.

Romesburg, C. H. (1984). *Cluster Analysis for Researchers.* Belmont, California: Lifetime Learning Publications.

Schneider, W., & Bjorklund, D. F. (1998). Memory. In D. Kuhn & R. S. Siegler (Eds.), *Handbook of child psychology: Cognition, perception, and language* (5th ed., pp. 467-521). New York: Wiley.

Schmidt, M. E. (2003). The relationship between the Wechsler Intelligence Scale for children – third edition and the Test of Memory and Learning in a pediatric Traumatic Brain Injury population. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 64*(3-B), 1518.

Schmitt, A. J., & Decker, S. L. (2008). Test Review of the TOMAL-2. *Journal of Psychoeducational Assessment.* Retrieved December 2nd, 2008 at http://jpa.sagepub.com/cgi/rapidpdf/0734282908322449v1.

Schneider, W., & Pressley, M. (1997). *Memory development between 2 and 20* (2nd Ed.) (pp. 89-120). New York: Springer-Verlag.

Seaton, B. E., Allen, D. N., Goldstein, G. G., Kelley, M. E., & van Kammen, D. P.

(1999). Relations between cognitive and symptom profile heterogeneity in schizophrenia. *The Journal of Nervous and Mental Disease, 187*(7), 414-419.

Sheslow, D., & Adams, W. (1990). *Wide range assessment of memory and learning.* Wilmington, DE: Jastak Associates.

Sheslow, D., & Adams, W. (2003). *Wide range assessment of memory and learning, (2nd ed.).* Psychological Assessment Resources, Inc.

Siegel, D. J. (2001). Memory: An overview, with emphasis on developmental, interpersonal, and neurobiological aspects. *Journal of the American Academy of Child and Adolescent Psychiatry, 40*(9), 997-1011.

Solokov, E. N. (1963). *Perception and the conditioned reflex.* Hillsdale, NJ: LEA.

Spiedel, G. E. (1991). Phonological short-term memory and individual differences in learning to speak: A bilingual case study. *First Language, 13,* 69-91.

Swanson, H.L. (1999) What develops in working memory? A life span perspective. *Developmental Psychology*, 35, 986–1000.

Tabares-Seisdedos, R., Balanza-Martinez, V., Moreno-Sanchez, J., Martinez-Aran, A., Salazar-Fraile, J., Selva-Vera, G., Rubio, C., Mata, I., Gomez-Beneyto, M., & Vieta, E. (2008). Neurocognitive and clinical predictors of functional outcome in patients with schizophrenia and bipolar I disorder at one-year follow up. *Journal of Affective Disorders, 109*, 286-299.

Temple, C. M., & Cornish, K. M. (1993). Recognition memory for words and faces in schoolchildren: A female advantage for words. *British Journal of Developmental Psychology, 11*(4), 421-426.

Thaler, N. S., Allen, D. N., McMurray, J. C., & Mayfield, J. (in press). Profiles of attention and memory deficits in children with ADHD assessed with the Test of Memory and Learning. *The Clinical Neuropsychologist.*

Thomason, M., Race, E., Burrows, B., Whitfield-Gabrieli, S., Glover, G., & Gabrieli, J. (2009). Development of spatial and verbal working memory capacity in the human brain. *Journal of Cognitive Neuroscience, 21*(2), 316-332.

Thompson, R. F. (2000). *The Brain: A Neuroscience Primer, 3rd edition.* New York, NY: Worth.

Tryon, R. C. (1939). *Cluster Analysis.* Edwards Brothers.

Tulving, E. (1984). Precise elements of episodic memory. *Behavioral and Brain Sciences, 7,* 223 – 268.

Tupler, L. A., Krishnan, K. R., Greenberg, D. L., Marcovina, S. M., Payne, M. E., MacFall, J. R., et al. (2007). Predicting memory decline in normal elderly: Genetics, MRI, and cognitive research. *Neurobiology of Aging, 28,* 1644-1656.

Ullman, D. G., McKee, D. T., Campbell, K. E., Larrabee, G. J., & Trahan, D. E. (1997). Preliminary children's norms for the Continuous Visual Memory Test. *Child Neuropsychology, 3*(3), 171-175.

Volkmar, F. R. & Greenough, W. T. (1972). Rearing complexity affects branching of dendrites in the visual cortex of the rat. *Science, 176*(4042), 1445-1447.

Voyer, D., Voyer, S., & Bryden, M. P. (1995). Magnitude of sex differences in spatial abilities: A meta-analysis and consideration of critical variables. *Psychological Bulletin, 117,* 250-270.

Ward, J. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association, 58,* 236-244.

Waxman, R. S., & Casey, J. E. (2006). Empirically derived ability-achievement subtypes in a clinical sample. *Child Neuropsychology, 12,* 23-38.

Wechsler, D. (1991). *Wechsler Intelligence Scale for Children-III.* New York: The Psychological Corporation.

Wechsler, D. (1995). *Children's Memory Scale*. San Antonio, TX: The Psychological Corporation.

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale – third edition.* San Antonio, TX: The Psychological Corporation.

Winston, J. (1993). The biology and function of rapid eye movement sleep. *Current Opinion in Neurobiology, 3,* 243-248.

Zouk, H., McGirr, A., Lebel, V., Benkelfat, C., Rouleau, G., & Turecki, G. (2007). The effect of genetic variation of the serotonin 1B receptor gene on impulsive aggressive behavior and suicide. *American Journal of Medical Genetics Part B, 144,* 996-1002.

Zhang, T., Ramakrishnon, R., and Livny M. (1996). BIRCH: An Efficient Data Clustering Method for Very Large Datebases. *Proceedings of the ACM SIGMOD Conference on Management of Data,* p. 103-114, Montreal, Canada.

APPENDIX A

TABLES

Table 1.

TOMAL and TOMAL-2 Index Subtest Composition

Core Verbal	Core Nonverbal	Delayed Recall*	Sequential Recall	Free Recall	Associative Recall	Learning	Attention/ Concentration
MFS	FM	MFSD	VSM	FM	MFS	WSR	DF
WSR	AVM	WSRD	DF	AVM	PR	VSR	DB
OR	VSM	FMD**	LF	MFL		OR	LF
PR	MFL	VSRD**	MI	OR**		PR	LB
DF**	VSR**						MI

*Note. Renamed “Verbal Delayed Recall” on the TOMAL-2
**Note. Not present in the TOMAL-2 version of this index

Table 2.

Four-Factor Solution of the TOMAL Standardization Sample (from Reynolds & Bigler, 1994)

Subtest	Factor 1	Factor 2	Factor 3	Factor 4
MFS	.37			
WSR	.73			
OR	.57			
PR	.68			
FM	.37			
VSR	.33			
DF		.74		
LF		.76		
MI		.38		
VSM		.38		
DB			.63	
LB			.66	
AVM				.39
MFL				.44

Note. MFS = Memory for Stories, FM = Facial Memory, WSR = Word Selective Reminding, VSR = Visual Selective Reminding, OR = Object Recall, AVM = Abstract Visual Memory, DF = Digits Forward, VSM = Visual Sequential Memory, PR = Paired Recall, MLA = Memory for Location, MI = Manual Imitation, LF = Letters Forward, DB = Digits Backward, LB = Letters Backward.

Table 3.

Cluster Analysis Methodology

Measure	Age	Attributes	Similarity Coefficient	Clustering Method	Clusters	Plotted Across
TOMAL	5-8 range	Scaled Subtest Scores	Squared Euclidean Distance	Ward's Minimum Variance	3, 4, 5, 6	VMI, NMI, ACI
	9-11 range					
	12-19 range Combined	Raw Subtest Scores*		Two-step Method**		SS*** Four Factors

**Note.* Raw subtest scores were only used to calculate clusters that differed by age.
***Note.* Two-step method provided meaningful data only with raw subtest scores on the entire standardization sample.
****Note.* SS = Standardization Sample

Table 4.

Four-cluster solution for the total sample using Ward's Method

TOMAL Subtest	C1 (n=314)		C2 (n=492)		C3 (n=169)		C4 (n=146)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
MFS	8.0	2.7	10.5	3.0	10.8	2.4	12.3	2.7
FM	8.4	2.3	10.0	2.5	11.5	2.6	11.8	3.2
WSR	8.3	2.5	10.5	2.5	11.7	2.7	12.0	2.8
VSR	8.3	3.2	9.5	2.8	12.1	2.7	12.2	3.2
OR	8.2	2.5	10.3	2.7	11.3	2.7	12.4	3.4
AVM	8.1	2.3	9.8	2.7	11.9	2.1	12.8	2.8
DF	7.6	2.6	10.6	2.9	9.3	2.1	13.7	3.0
VSM	8.4	2.0	9.7	2.8	11.8	2.4	11.8	3.8
PR	7.8	2.5	10.4	2.5	11.4	2.3	11.3	2.7
MLA	8.3	3.3	8.6	2.9	13.4	3.7	13.8	2.9
MI	7.6	2.6	9.9	2.4	10.1	2.7	13.7	3.2
LF	7.6	2.3	10.3	2.7	9.8	2.1	13.3	2.8
DB	8.1	2.5	11.0	3.2	10.6	2.4	14.4	3.1
LB	7.9	2.5	10.6	3.0	10.2	2.9	14.3	3.0

Note. MFS = Memory for Stories, FM = Facial Memory, WSR = Word Selective Reminding, VSR = Visual Selective Reminding, OR = Object Recall, AVM = Abstract Visual Memory, DF = Digits Forward, VSM = Visual Sequential Memory, PR = Paired Recall, MLA = Memory for Location, MI = Manual Imitation, LF = Letters Forward, DB = Digits Backward, LB = Letters Backward.

Table 5.

Five-cluster solution for the total sample using Ward's Method

TOMAL Subtest	C1 (n=314)		C2 (n=215)		C3 (n=277)		C4 (n=169)		C5 (n=146)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
MFS	8.0	2.7	10.8	3.4	10.8	2.4	10.0	2.4	12.3	2.7
FM	8.5	2.3	10.6	2.3	11.4	2.6	9.1	2.4	11.8	3.2
WSR	8.3	2.5	10.1	2.4	11.7	2.7	10.9	2.5	12.0	2.8
VSR	8.3	3.2	9.5	3.0	12.1	2.7	9.6	2.5	12.1	3.2
OR	8.2	2.5	10.1	2.5	11.3	2.7	10.8	2.9	12.4	3.4
AVM	8.1	2.3	9.5	2.8	11.9	2.1	10.2	2.3	12.8	2.8
DF	7.6	2.6	9.5	2.6	9.3	2.1	12.1	2.7	13.7	3.0
VSM	8.4	2.0	9.2	2.9	11.8	2.4	10.4	2.3	11.8	3.8
PR	7.8	2.5	10.3	2.6	11.4	2.3	10.6	2.1	11.3	2.7
MLA	8.3	3.3	8.1	2.9	13.4	3.7	9.3	2.8	13.8	2.9
MI	7.6	2.6	9.8	2.3	10.1	2.7	10.0	2.6	13.7	3.2
LF	7.6	2.3	9.5	2.4	9.8	2.1	11.5	2.6	13.3	2.8
DB	8.1	2.5	9.9	2.5	10.6	2.4	12.5	3.5	14.4	3.1
LB	7.9	2.5	9.4	2.5	10.2	2.9	12.4	2.7	14.3	3.0

Note. See Table 3 for explanation of abbreviations.

Table 6.

Six-cluster solution for the total sample using Ward's Method

TOMAL Subtest	C1 (n=314)		C2 (n=215		C3 (n=277)		C4 (n=106)		C5 (n=63)		C6 (n=146)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
MFS	8.0	2.7	10.8	3.4	10.8	2.4	10.0	2.4	11.7	2.5	13.3	2.7
FM	8.5	2.3	10.6	2.3	11.5	2.6	9.1	2.4	11.3	3.2	12.8	3.1
WSR	8.3	2.5	10.1	2.4	11.7	2.7	10.9	2.5	11.8	3.0	12.2	2.6
VSR	8.3	3.2	9.5	3.0	12.1	2.7	9.6	2.5	10.9	2.8	14.1	2.7
OR	8.2	2.5	10.1	2.5	11.3	2.7	10.8	2.9	11.4	2.8	14.1	3.6
AVM	8.1	2.3	9.5	2.8	11.9	2.1	10.2	2.3	13.0	2.5	12.4	3.3
DF	7.6	2.6	9.5	2.6	9.3	2.1	12.1	2.7	12.2	2.3	16.1	2.3
VSM	8.4	2.0	9.2	2.9	11.8	2.4	10.4	2.3	10.3	3.4	14.0	3.2
PR	7.8	2.5	10.3	2.6	11.4	2.3	10.6	2.1	11.2	2.7	11.5	2.6
MLA	8.3	3.3	8.1	2.9	13.4	3.7	9.3	2.8	14.1	2.8	13.5	3.1
MI	7.6	2.6	9.8	2.3	10.1	2.7	10.0	2.6	12.2	2.6	16.1	2.4
LF	7.6	2.3	9.5	2.4	9.8	2.1	11.5	2.6	12.0	2.2	15.5	2.2
DB	8.1	2.5	9.9	2.5	10.6	2.4	12.5	3.4	14.9	2.9	13.7	3.4
LB	7.9	2.5	9.4	2.5	10.2	2.9	12.4	2.7	14.4	2.5	14.3	3.7

Note. See Table 3 for explanation of abbreviations.

Table 7.

Cross-Tabulation for Ward's Method and K-mean's Iterations: 4 Cluster Solution.

		K-means Iteration				
Ward's Method		1	2	3	4	Total
1	Count	252	18	5	0	275
	Agreement	91.6%	6.5%	1.8%	.0%	100.0%
2	Count	77	346	62	42	527
	Agreement	14.6%	65.7%	11.8%	8.0%	100.0%
3	Count	0	11	139	6	156
	Agreement	.0%	7.1%	89.1%	3.8%	100.0%
4	Count	1	7	23	132	163
	Agreement	29.4%	34.1%	20.4%	16.1%	100.0%

Note. Kappa = .69, n = 1121, T = 38.84, $p < .001$.

Table 8.

Cross-Tabulation for Ward's Method and K-mean's Iterations: 5 Cluster Solution.

		K-means Iteration					
Ward's Method		1	2	3	4	5	Total
1	Count	233	37	3	2	0	275
	Agreement	84.7%	13.5%	1.1%	.7%	.0%	100.0%
2	Count	37	208	19	49	1	314
	Agreement	11.8%	66.2%	6.1%	15.6%	.3%	100.0%
3	Count	0	18	133	0	5	156
	Agreement	.0%	11.5%	85.3%	.0%	3.2%	100.0%
4	Count	4	29	28	148	4	213
	Agreement	1.9%	13.6%	13.1%	69.5%	1.9%	100.0%
5	Count	1	0	29	27	106	163
	Agreement	.6%	.0%	17.8%	16.6%	65.0%	100.0%

Note. Kappa = .67, n = 1121, T = 43.77, $p < .001$.

Table 9.

Cross-Tabulation for Ward's Method and K-mean's Iterations: 6 Cluster Solution.

		K-means Iteration						
Ward's Method		1	2	3	4	5	6	Total
1	Count	221	42	7	5	0	0	275
	Agreement	80.4%	15.3%	2.5%	1.8%	.0%	.0%	100.0%
2	Count	35	197	20	51	11	0	314
	Agreement	11.1%	62.7%	6.4%	16.2%	3.5%	.0%	100.0%
3	Count	0	16	118	0	22	0	156
	Agreement	.0%	10.3%	75.6%	.0%	14.1%	.0%	100.0%
4	Count	3	23	20	145	18	4	213
	Agreement	1.4%	10.8%	9.4%	68.1%	8.5%	1.9%	100.0%
5	Count	1	0	3	7	84	5	100
	Agreement	1.0%	.0%	3.0%	7.0%	84.0%	5.0%	100.0%
6	Count	0	0	4	3	5	51	100
	Agreement	.0%	.0%	6.3%	4.8%	7.9%	81.0%	100.0%

Note. Kappa = .66, n = 1121, T = 45.79, $p < .001$.

Table 10.

Three-cluster solution for the total sample using the Two-Cluster Method

TOMAL Subtest	C1		C2		C3	
	Mean	SD	Mean	SD	Mean	SD
MFS	9.5	3.1	9.8	3.1	11.6	3.0
FM	9.5	2.6	9.9	2.7	11.1	2.0
WSR	9.2	2.9	10.2	2.7	11.9	2.5
VSR	8.7	3.0	10.0	3.1	11.2	3.3
OR	9.0	2.8	10.3	2.8	11.7	3.3
AVM	9.1	2.5	10.1	3.0	11.3	3.0
DF	8.9	3.0	9.6	3.0	12.3	3.3
VSM	9.3	2.6	9.7	2.8	11.2	3.3
PR	8.8	2.9	10.1	2.6	11.4	2.3
MLA	8.9	2.7	9.6	4.2	11.7	4.0
MI	9.0	2.6	9.1	2.9	12.2	3.2
LF	8.8	2.8	9.6	2.8	12.1	2.8
DB	10.4	3.8	9.7	3.0	12.7	3.1
LB	9.1	2.6	9.6	3.1	13.3	3.1

Note. MFS = Memory for Stories, FM = Facial Memory, WSR = Word Selective Reminding, VSR = Visual Selective Reminding, OR = Object Recall, AVM = Abstract Visual Memory, DF = Digits Forward, VSM = Visual Sequential Memory, PR = Paired Recall, MLA = Memory for Location, MI = Manual Imitation, LF = Letters Forward, DB = Digits Backward, LB = Letters Backward.

Table 11.

Cross-Tabulation for Two-Step Method and K-mean's Iterations: 3 Cluster Solution.

		K-means Iteration			
Two-Step Method		1	2	3	Total
1	Count	327	46	0	373
	Agreement	87.7%	12.3%	.0%	100.0%
2	Count	14	428	39	481
	Agreement	2.9%	89.0%	8.1%	100.0%
3	Count	0	10	257	267
	Agreement	.0%	3.7%	97.3%	100.0%

Note. Kappa = .85, n = 1121, T = 39.9, $p < .001$.

Table 12.

Age Differences of the 5-Cluster Solution

Cluster	Mean Age	SD	Scheffe*
Low Functioning (C1)	10.5	3.6	C1<C5
Average (C2)	9.8	3.1	C2<C4, C5
High Nonverbal (C3)	10.2	3.5	C3<C5
High Attention (C4)	11.0	3.6	C4>C2
High Functioning (C5)	11.6	3.4	C5>C3, C2, C1

*Note. $p < .05$.

Table 13.

Four-cluster solution for the 5-8 age sample using Ward's Method (N=361)

TOMAL Subtest	C1 (n=78)		C2 (n=72)		C3 (n=147)		C4 (n=64)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
MFS	8.1	2.8	9.1	2.2	11.0	2.7	11.8	3.0
FM	8.2	3.1	10.0	2.0	10.1	2.4	11.7	3.3
WSR	7.8	3.2	9.8	2.3	10.4	2.4	11.8	2.5
VSR	6.6	2.8	9.3	2.0	10.1	2.4	11.8	2.1
OR	7.6	2.1	9.0	1.9	10.4	2.6	13.1	2.9
AVM	7.6	2.0	9.2	2.0	10.7	2.4	12.7	2.9
DF	7.8	2.3	8.1	2.3	10.5	2.7	13.0	2.8
VSM	8.4	2.0	9.6	2.4	10.1	2.5	11.9	3.6
PR	7.3	2.7	9.3	2.0	10.3	2.1	11.4	2.0
MLA	7.6	2.2	9.5	1.6	10.0	2.7	12.8	2.8
MI	8.6	2.8	8.8	2.2	10.1	2.2	12.3	2.7
LF	7.9	2.2	7.9	2.2	10.4	2.5	13.2	3.2
DB	8.9	3.6	10.6	3.7	12.1	3.5	12.8	3.1
LB	7.7	2.4	9.1	1.8	10.6	2.6	13.4	3.5

Note. MFS = Memory for Stories, FM = Facial Memory, WSR = Word Selective Reminding, VSR = Visual Selective Reminding, OR = Object Recall, AVM = Abstract Visual Memory, DF = Digits Forward, VSM = Visual Sequential Memory, PR = Paired Recall, MLA = Memory for Location, MI = Manual Imitation, LF = Letters Forward, DB = Digits Backward, LB = Letters Backward.

Table 14.

Cross-Tabulation for Ward's Method and K-mean's Iterations of 5-8 Age Group: 4 Cluster Solution.

		K-means Iteration				
Ward's Method		1	2	3	4	Total
1	Count	64	13	1	0	78
	Agreement	82.1%	16.7%	0.1%	.0%	100.0%
2	Count	3	61	8	0	72
	Agreement	0.4%	84.7%	11.1%	.0%	100.0%
3	Count	0	40	102	5	147
	Agreement	.0%	27.2%	69.4%	3.4%	100.0%
4	Count	0	1	16	47	64
	Agreement	.0%	.2%	25.0%	73.4%	100.0%

Note. Kappa = .67, n = 361, T = 21.75, $p < .001$.

Table 15.

Cross-Tabulation for Ward's Method and K-mean's Iterations of 5-8 Age Group: 3 Cluster Solution.

		K-means Iteration			
Two-Step Method		1	2	3	Total
1	Count	53	2	0	55
	Agreement	96.4%	3.6%	.0%	100.0%
2	Count	29	135	14	178
	Agreement	16.3%	75.8%	7.9%	100.0%
3	Count	0	17	111	128
	Agreement	.0%	13.3%	86.7%	100.0%

Note. Kappa = .73, n = 361, T = 19.08, $p < .001$.

Table 16.

Four-cluster solution for the 9-11 age sample using Ward's Method (n=377)

TOMAL Subtest	C1 (n=171)		C2 (n=146)		C3 (n=35)		C4 (n=25)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
MFS	9.2	2.8	10.8	2.6	10.7	2.2	13.0	3.1
FM	9.6	2.4	10.8	2.8	10.2	1.8	13.2	2.9
WSR	8.6	2.2	12.1	2.6	10.0	1.6	13.5	3.5
VSR	9.2	3.4	10.0	3.0	12.1	3.0	13.3	3.5
OR	9.1	2.6	11.5	2.6	9.5	2.7	14.8	3.6
AVM	9.0	2.6	11.2	2.6	12.0	2.6	14.8	2.6
DF	8.7	2.9	11.3	3.3	10.6	1.9	13.6	3.6
VSM	9.0	2.3	10.9	2.6	10.8	2.7	13.6	4.1
PR	9.2	3.0	11.5	2.5	9.6	2.4	12.1	2.7
MLA	7.8	2.9	8.8	2.8	17.9	1.9	15.4	3.5
MI	8.1	2.3	10.0	2.9	9.9	2.4	14.4	3.7
LF	8.7	2.4	11.0	2.7	11.0	2.2	12.7	2.5
DB	9.0	2.3	11.8	3.0	10.9	2.8	13.9	3.0
LB	8.6	2.3	11.4	3.2	10.0	2.7	15.2	1.7

Note. MFS = Memory for Stories, FM = Facial Memory, WSR = Word Selective Reminding, VSR = Visual Selective Reminding, OR = Object Recall, AVM = Abstract Visual Memory, DF = Digits Forward, VSM = Visual Sequential Memory, PR = Paired Recall, MLA = Memory for Location, MI = Manual Imitation, LF = Letters Forward, DB = Digits Backward, LB = Letters Backward.

Table 17.

Five-cluster solution for the 9-11 age sample using Ward's Method (n=377)

TOMAL Subtest	C1 (n=171)		C2 (n=105)		C3 (n=41)		C4 (n=35)		C5 (n=25)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
MFS	9.2	2.8	10.1	2.1	12.6	2.9	10.7	2.2	13.0	3.1
FM	9.6	2.4	9.9	2.4	13.1	2.7	10.2	1.8	13.2	2.9
WSR	8.6	2.2	12.0	2.7	12.6	2.3	10.0	1.6	13.5	3.5
VSR	9.2	3.4	9.2	2.6	12.0	3.0	12.1	3.0	13.3	3.5
OR	9.1	2.6	11.4	2.8	11.7	2.2	9.5	2.7	14.8	3.6
AVM	9.0	2.6	10.7	2.5	12.4	2.2	12.0	2.6	14.8	2.6
DF	8.7	2.9	12.2	3.1	8.9	2.3	10.6	1.9	13.6	3.6
VSM	9.0	2.3	10.6	2.5	11.7	2.8	10.8	2.7	13.6	4.1
PR	9.2	3.0	11.1	2.3	12.3	2.7	9.6	2.4	12.1	2.7
MLA	7.8	2.9	8.6	2.8	9.4	2.7	17.9	1.9	15.4	3.5
MI	8.1	2.3	10.0	2.9	9.8	3.2	9.9	2.4	14.4	3.7
LF	8.7	2.4	11.5	2.7	9.8	2.2	11.0	2.2	12.7	2.5
DB	9.0	2.3	12.8	2.7	9.2	2.3	10.9	2.8	13.9	3.0
LB	8.6	2.3	12.4	2.9	8.8	2.6	10.0	2.7	15.2	1.7

Note. See Table 16 for explanation of abbreviations.

Table 18.

Six-cluster solution for the 9-11 sample using Ward's Method (n=377)

TOMAL Subtest	C1 (n=70)		C2 (n=101		C3 (n=105)		C4 (n=41)		C5 (n=35)		C6 (n=25)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
MFS	10.1	2.5	8.5	2.9	10.1	2.1	12.6	2.9	10.7	2.2	13.0	3.1
FM	9.7	2.7	9.6	2.1	9.9	2.4	13.1	2.7	10.2	1.8	13.2	2.9
WSR	7.8	1.9	9.2	2.2	12.0	2.7	12.6	2.3	10.0	1.6	13.5	3.5
VSR	6.6	2.2	11.1	2.9	9.2	2.6	12.0	3.0	12.1	3.0	13.3	3.5
OR	8.3	2.2	9.6	2.7	11.4	2.8	11.7	2.2	9.5	2.7	14.8	3.6
AVM	8.2	2.3	9.6	2.7	10.7	2.5	12.4	2.2	12.0	2.6	14.8	2.6
DF	8.2	3.2	8.9	2.7	12.2	3.1	8.9	2.3	10.6	1.9	13.6	3.6
VSM	8.4	2.2	9.3	2.3	10.6	2.5	11.7	2.8	10.8	2.7	13.6	4.1
PR	7.9	2.7	10.0	3.0	11.1	2.3	12.3	2.7	9.6	2.4	12.1	2.7
MLA	6.9	2.8	8.4	2.9	8.6	2.8	9.4	2.7	17.9	1.9	15.4	3.5
MI	7.8	2.7	8.2	2.0	10.0	2.9	9.8	3.2	9.9	2.4	14.4	3.7
LF	8.9	3.0	8.6	2.0	11.5	2.7	9.8	2.2	11.0	2.2	12.7	2.5
DB	8.7	2.6	9.1	2.1	12.8	2.7	9.2	2.3	10.9	2.8	13.9	3.0
LB	8.4	2.3	8.8	2.3	12.4	2.8	8.8	2.6	10.0	2.7	15.2	1.7

Note. See Table 16 for explanation of abbreviations.

Table 19.

Cross-Tabulation for Ward's Method and K-mean's Iterations for 9-11 age group: 4 Cluster Solution.

		K-means Iteration				
Ward's Method		1	2	3	4	Total
1	Count	133	17	21	0	171
	Agreement	77.8%	9.9%	12.3%	.0%	100.0%
2	Count	14	104	23	5	146
	Agreement	9.6%	71.2%	15.8%	3.4%	100.0%
3	Count	0	1	32	2	35
	Agreement	.0%	2.9%	91.4%	5.7%	100.0%
4	Count	0	0	0	25	25
	Agreement	.0%	.0%	.0%	100.0%	100.0%

Note. Kappa = .67, n = 377, T = 20.27, $p < .001$.

Table 20.

Cross-Tabulation for Ward's Method and K-mean's Iterations for 9-11 age group: 5 Cluster Solution.

		K-means Iteration					
Ward's Method		1	2	3	4	5	Total
1	Count	114	8	42	7	0	171
	Agreement	66.7%	4.7%	24.6%	4.1%	.0%	100.0%
2	Count	6	73	20	3	3	105
	Agreement	5.7%	69.5%	19.0%	2.9%	2.9%	100.0%
3	Count	0	1	39	1	0	41
	Agreement	.0%	2.4%	95.1%	2.4%	.0%	100.0%
4	Count	0	1	0	32	2	35
	Agreement	.0%	2.9%	.0%	91.4%	5.7%	100.0%
5	Count	0	1	2	0	22	25
	Agreement	.0%	4.0%	8.0%	.0%	88.0%	100.0%

Note. Kappa = .66, n = 377, T = 24.00, $p < .001$.

Table 21.

Cross-Tabulation for Ward's Method and K-mean's Iterations 9-11 age group: 6 Cluster Solution.

		K-means Iteration						
Ward's Method		1	2	3	4	5	6	Total
1	Count	53	13	0	4	0	0	70
	Agreement	75.7%	18.6%	.0%	5.7%	.0%	.0%	100.0%
2	Count	40	40	16	5	0	0	101
	Agreement	39.6%	39.6%	15.8%	5.0%	.0%	.0%	100.0%
3	Count	4	24	3	59	15	0	105
	Agreement	3.8%	22.9%	2.9%	56.2%	14.3%	.0%	100.0%
4	Count	0	20	16	1	4	0	41
	Agreement	.0%	48.8%	39.0%	2.4%	9.8%	.0%	100.0%
5	Count	0	0	29	0	0	6	35
	Agreement	.0%	.0%	82.9%	.0%	17.1%	.0%	100.0%
6	Count	0	0	0	0	19	6	25
	Agreement	.0%	.0%	.0%	.0%	76.0%	24.0%	100.0%

Note. Kappa = .12, n = 377, T = 4.72, $p < .001$.

Table 22.

Gender differences per cluster on the 9-11 age group.

		Ward Method					
		C1	*C2*	*C3*	*C4*	*C5*	*Total*
Male	*Count*	*85*	*39*	*8*	*23*	*16*	*171*
	% within gender	*49.7%*	*22.8%*	*4.7%*	*13.5%*	*9.4%*	*100.0%*
Female	*Count*	*84*	*59*	*28*	*11*	*9*	*191*
	% within gender	*44.0%*	*30.9%*	*14.7%*	*5.8%*	*4.7%*	*100.0%*
Total	*Count*	*169*	*98*	*36*	*34*	*25*	*362*
	% within gender	*46.7%*	*27.1%*	*9.9%*	*9.4%*	*6.9%*	*100.0%*

Table 23.

Four-cluster solution for the 12-19 age sample using Ward's Method (N=383)

TOMAL Subtest	C1 (n=101)		C2 (n=142)		C3 (n=106)		C4 (n=34)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
MFS	7.0	2.7	10.4	2.5	12.1	3.3	13.1	2.2
FM	8.3	2.5	9.3	2.5	11.6	2.4	12.6	3.1
WSR	8.8	2.6	10.4	2.1	11.8	2.4	12.7	2.2
VSR	9.5	2.8	9.5	2.9	10.6	3.4	15.0	1.9
OR	8.8	2.9	10.1	2.6	11.1	2.5	13.9	2.8
AVM	8.2	2.6	9.2	3.1	11.5	2.3	11.8	1.9
DF	7.4	2.8	10.5	2.3	10.8	2.7	16.4	2.7
VSM	8.0	2.1	9.6	2.4	10.2	3.9	13.8	2.4
PR	8.2	2.8	10.3	2.3	12.0	2.1	11.0	2.9
MLA	8.6	4.5	8.7	3.7	13.5	2.7	12.8	2.9
MI	7.5	2.8	10.1	2.4	11.5	2.5	16.9	2.1
LF	7.6	2.4	10.4	2.2	10.4	2.9	16.0	1.3
DB	7.5	2.5	9.7	2.4	12.8	2.9	14.1	3.8
LB	7.4	2.6	10.7	3.2	13.1	2.5	15.2	3.6

Note. MFS = Memory for Stories, FM = Facial Memory, WSR = Word Selective Reminding, VSR = Visual Selective Reminding, OR = Object Recall, AVM = Abstract Visual Memory, DF = Digits Forward, VSM = Visual Sequential Memory, PR = Paired Recall, MLA = Memory for Location, MI = Manual Imitation, LF = Letters Forward, DB = Digits Backward, LB = Letters Backward.

Table 24.

Five-cluster solution for the 12-19 age sample using Ward's Method (N=383)

TOMAL Subtest	C1 (n=101)		C2 (n=142)		C3 (n=92)		C4 (n=34)		C5 (n=14)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
MFS	7.0	2.7	10.4	2.5	11.1	2.2	13.1	2.2	18.6	2.3
FM	8.3	2.5	9.3	2.5	11.7	2.5	12.6	3.1	10.7	1.2
WSR	8.8	2.6	10.4	2.1	12.1	2.4	12.7	2.2	9.5	.8
VSR	9.5	2.8	9.5	2.9	11.0	2.9	15.0	1.9	8.0	5.3
OR	8.8	2.9	10.1	2.6	11.3	2.3	13.9	2.8	9.8	3.6
AVM	8.2	2.6	9.2	3.1	11.5	2.3	11.8	1.9	11.6	1.7
DF	7.4	2.8	10.5	2.3	10.8	2.7	16.4	2.7	10.6	2.3
VSM	8.0	2.1	9.6	2.4	11.4	2.2	13.8	2.4	2.0	2.4
PR	8.2	2.8	10.3	2.3	12.3	1.9	11.0	2.9	9.9	2.4
MLA	8.6	4.5	8.7	3.7	13.2	2.8	12.8	2.9	15.2	1.1
MI	7.5	2.8	10.1	2.4	11.4	2.6	16.9	2.1	12.1	1.0
LF	7.6	2.4	10.4	2.2	10.6	2.9	16.0	1.3	8.9	2.2
DB	7.5	2.5	9.7	2.4	13.0	2.8	14.1	3.8	11.9	3.4
LB	7.4	2.6	10.7	3.2	13.3	2.1	15.2	3.6	10.9	4.0

Note. See Table 23 for explanation of abbreviations.

Table 25.

Six-cluster solution for the 12-9 age sample using Ward's Method (n=383)

TOMAL Subtest	C1 (n=73)		C2 (n=28)		C3 (n=142)		C4 (n=92)		C5 (n=34)		C6 (n=14)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
MFS	7.5	2.8	5.9	2.0	10.4	2.5	11.1	2.2	13.1	2.2	18.6	2.3
FM	8.1	2.6	8.9	2.1	9.3	2.5	11.7	2.5	12.6	3.1	10.7	1.2
WSR	9.2	2.6	7.8	2.3	10.4	2.1	12.1	2.4	12.7	2.2	9.6	.8
VSR	10.2	2.6	7.5	2.5	9.5	2.9	11.0	2.9	15.0	1.9	8.0	5.3
OR	9.2	2.5	7.5	3.6	10.1	2.6	11.3	2.3	13.9	2.8	9.8	3.6
AVM	8.3	2.6	8.0	2.6	9.2	3.1	11.5	2.3	11.8	1.9	11.6	1.7
DF	7.3	2.7	7.8	2.9	10.5	2.3	10.8	2.7	16.4	2.7	10.6	2.3
VSM	8.2	2.1	7.5	2.1	9.6	2.4	11.4	2.2	13.8	2.4	2.0	2.4
PR	8.8	2.4	6.6	3.2	10.3	2.3	12.3	1.9	11.0	2.9	9.9	2.4
MLA	10.7	3.1	3.1	2.7	8.7	3.7	13.2	2.8	12.8	2.9	15.2	1.1
MI	8.0	2.6	6.0	2.7	10.1	2.4	11.4	2.6	16.9	2.1	12.1	1.0
LF	7.5	2.3	7.8	2.8	10.4	2.2	10.6	2.9	16.0	1.3	8.9	2.2
DB	7.4	2.7	7.8	1.7	9.7	2.4	13.0	2.8	14.1	3.8	11.9	3.4
LB	7.8	2.5	6.4	2.5	10.7	3.2	13.3	2.1	15.2	3.6	10.9	4.0

Note. See Table 23 for explanation of abbreviations.

Table 26.

Cross-Tabulation for Ward's Method and K-mean's Iterations for 12-19 age group: 4 Cluster Solution.

		K-means Iteration				
Ward's Method		1	2	3	4	Total
1	Count	89	9	0	0	98
	Agreement	90.8%	9.2%	.0%	.0%	100.0%
2	Count	9	114	7	0	130
	Agreement	6.9%	87.7%	5.4%	.0%	100.0%
3	Count	3	19	93	0	115
	Agreement	2.6%	16.5%	80.9%	.0%	100.0%
4	Count	0	0	6	34	40
	Agreement	.0%	.0%	15.0%	85.0%	100.0%

Note. Kappa = .81, n = 383, T = 25.65, $p < .001$.

Table 27.

Cross-Tabulation for Ward's Method and K-mean's Iterations for 9-11 age group: 5 Cluster Solution.

		K-means Iteration					
Ward's Method		1	2	3	4	5	Total
1	Count	90	12	0	0	0	102
	Agreement	88.2%	11.8%	.0%	.0%	.0%	100.0%
2	Count	5	110	3	0	0	118
	Agreement	4.2%	93.2%	2.5%	.0%	.0%	100.0%
3	Count	5	20	83	0	0	108
	Agreement	4.6%	18.5%	76.9%	.0%	.0%	100.0%
4	Count	0	0	6	34	0	40
	Agreement	.0%	.0%	15.0%	85.0%	.0%	100.0%
5	Count	1	0	0	0	14	15
	Agreement	6.7%	.0%	.0%	.0%	93.3%	100.0%

Note. Kappa = .82, n = 383, T = 27.74, $p < .001$.

Table 28.

Cross-Tabulation for Ward's Method and K-mean's Iterations 12-19 age group: 6 Cluster Solution.

		K-means Iteration						
Ward's Method		1	2	3	4	5	6	Total
1	Count	59	0	12	0	0	0	71
	Agreement	83.1%	.0%	16.9%	.0%	.0%	.0%	100.0%
2	Count	7	28	4	0	0	0	39
	Agreement	17.9%	71.8%	10.3%	.0%	.0%	.0%	100.0%
3	Count	3	24	107	3	0	0	113
	Agreement	2.7%	.0%	94.7%	2.7%	.0%	.0%	100.0%
4	Count	3	19	83	0	0	0	105
	Agreement	2.9%	.0%	18.1%	79.0%	.0%	.0%	100.0%
5	Count	0	0	0	6	34	0	40
	Agreement	.0%	.0%	.0%	15.0%	85.0%	.0%	100.0%
6	Count	1	0	0	0	0	14	15
	Agreement	6.7%	.0%	.0%	.0%	93.3%	24.0%	100.0%

Note. Kappa = .80, n = 383, T = 30.53, $p < .001$.

Table 29.

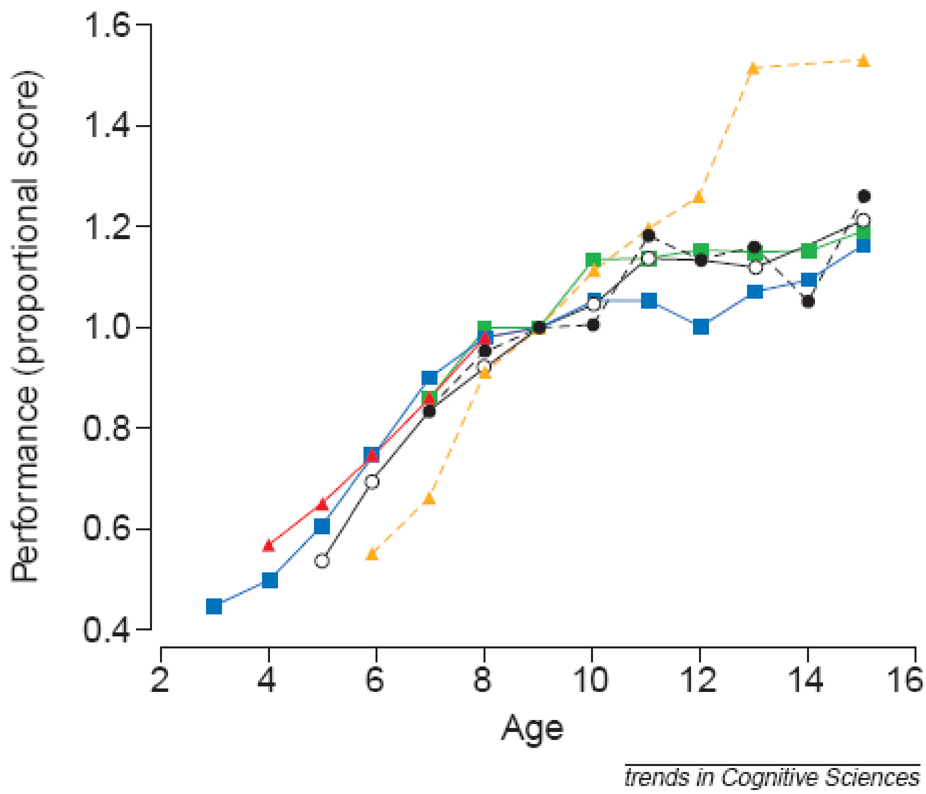
Gender differences per cluster on the 12-19 age group.

			F	M	Total
1	Count	0	36	37	73
	% within Ward Method	.0%	49.3%	50.7%	100.0%
2	Count	0	15	13	28
	% within Ward Method	.0%	53.6%	46.4%	100.0%
3	Count	0	63	79	142
	% within Ward Method	.0%	44.4%	55.6%	100.0%
4	Count	4	55	33	92
	% within Ward Method	4.3%	59.8%	35.9%	100.0%
5	Count	0	20	14	34
	% within Ward Method	.0%	58.8%	41.2%	100.0%
6	Count	0	6	8	14
	% within Ward Method	.0%	42.9%	57.1%	100.0%
Total	Count	4	195	184	383
	% within Ward Method	1.0%	50.9%	48.0%	100.0%

APPENDIX B

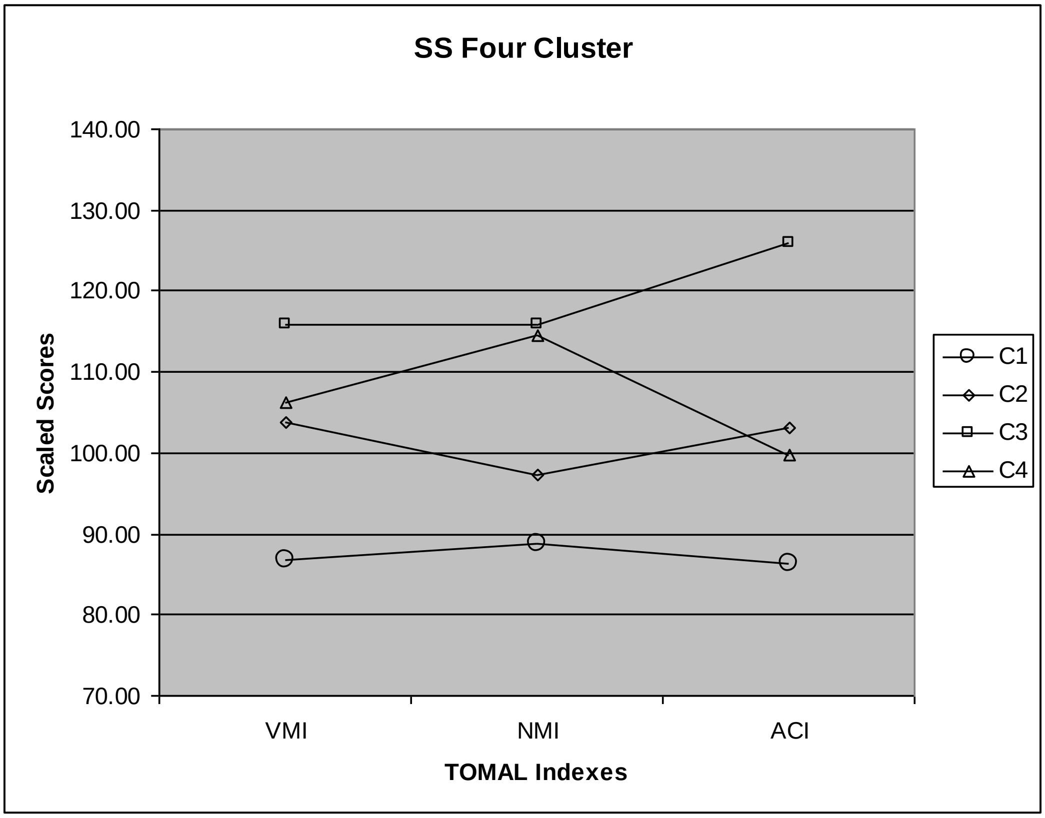
FIGURES

Figure 1. *Performance on measures of short-term memory as a function of age. Mean performance of each age group is plotted as a proportion of mean performance of nine-year olds. (from Gathercole, 1999).*



Note. *Blue squares, digit span (phonological memory); red triangles, non-word repetition (phonological memory); open circles, forward digit span; green squares, Corsi blocks (visuospatial memory); yellow triangles, listening span (complex working memory); filled circles, backward digit span (complex working memory).*

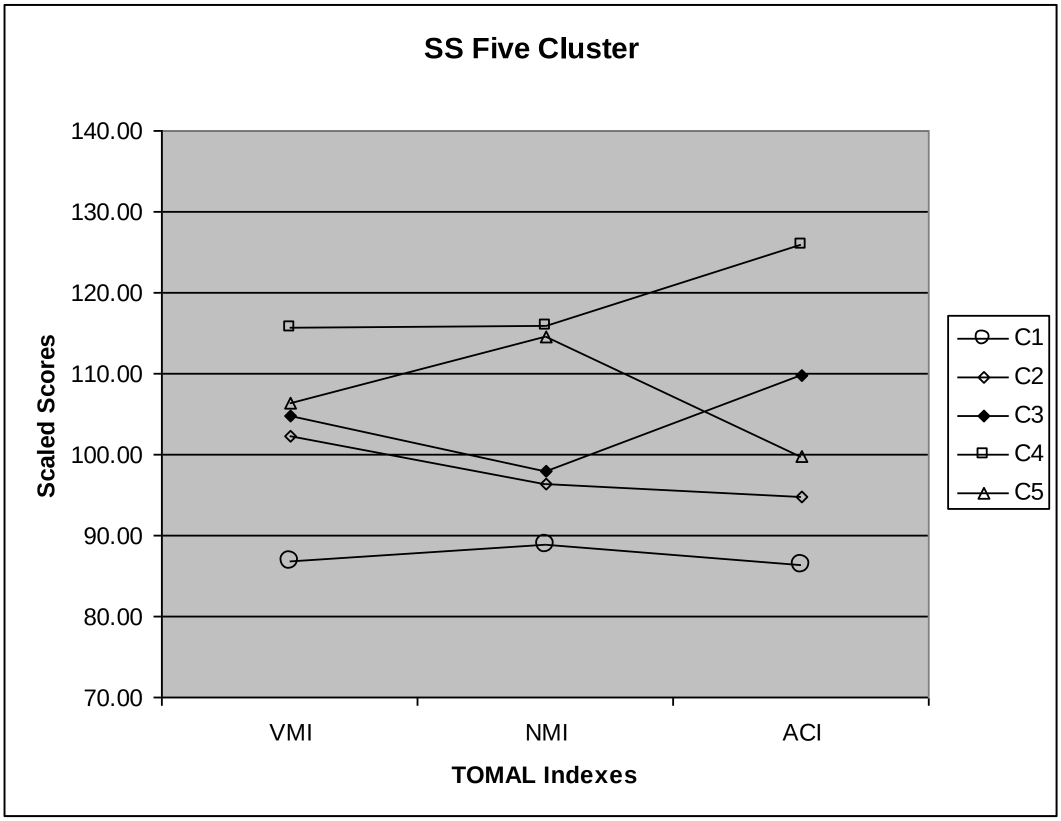
Figure 2. Cluster Profiles of TOMAL Indexes for Four Cluster Solutions: Scaled Scores and Ward's Method.



Note. VMI = Verbal Memory Index, NMI = Nonverbal Memory Index, ACI = Attention/Concentration Index.

Note. C1 = Low Average, C2 = Average, C3 = High Average, C4 = Nonverbal

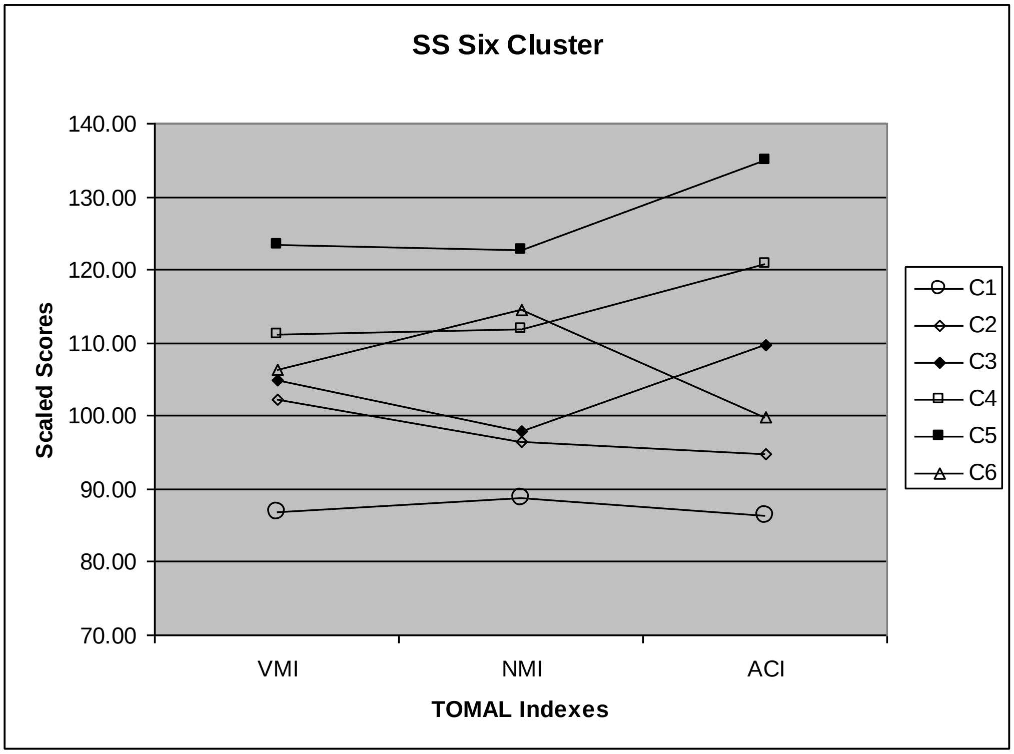
Figure 3. Cluster Profiles of TOMAL Indexes for the Five Cluster Solution: Scaled Scores and Ward's Method.



Note. VMI = Verbal Memory Index, NMI = Nonverbal Memory Index, ACI = Attention/Concentration Index.

Note. C1 = Low Average, C2 = Average, C3 = Verbal, C4 = High Average, C5 = Nonverbal

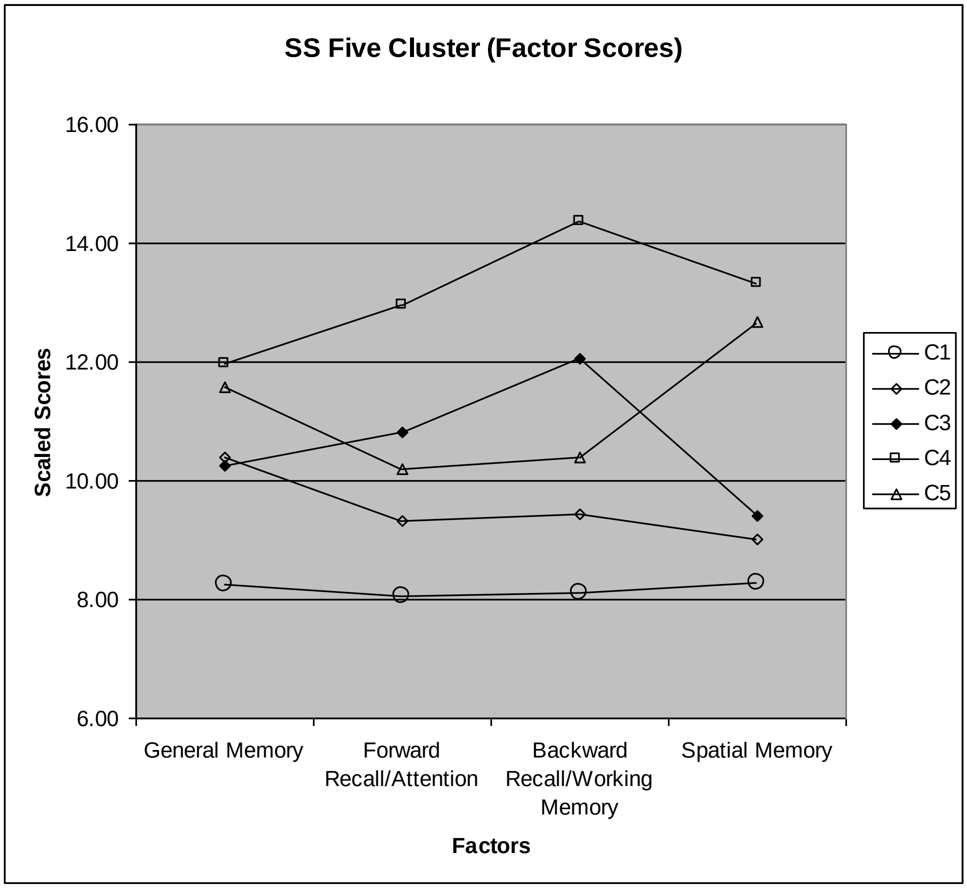
Figure 4. Cluster Profiles of TOMAL Indexes for the Six Cluster Solution: Scaled Scores and Ward's Method.



Note. VMI = Verbal Memory Index, NMI = Nonverbal Memory Index, ACI = Attention/Concentration Index.

Note. C1 = Low Average, C2 = Average, C3 = Verbal, C4 = High Average, C5 = Superior, C6 = Nonverbal

Figure 5. Cluster Profiles of Factor Scores for the Five Cluster Solutions: Scaled Scores and Ward's Method.



Note. C1 = Low Average, C2 = Average, C3 = Verbal, C4 = High Average, C5 = Nonverbal

Figure 6. 4-Cluster DFA of Scaled Scores Using Ward's Method

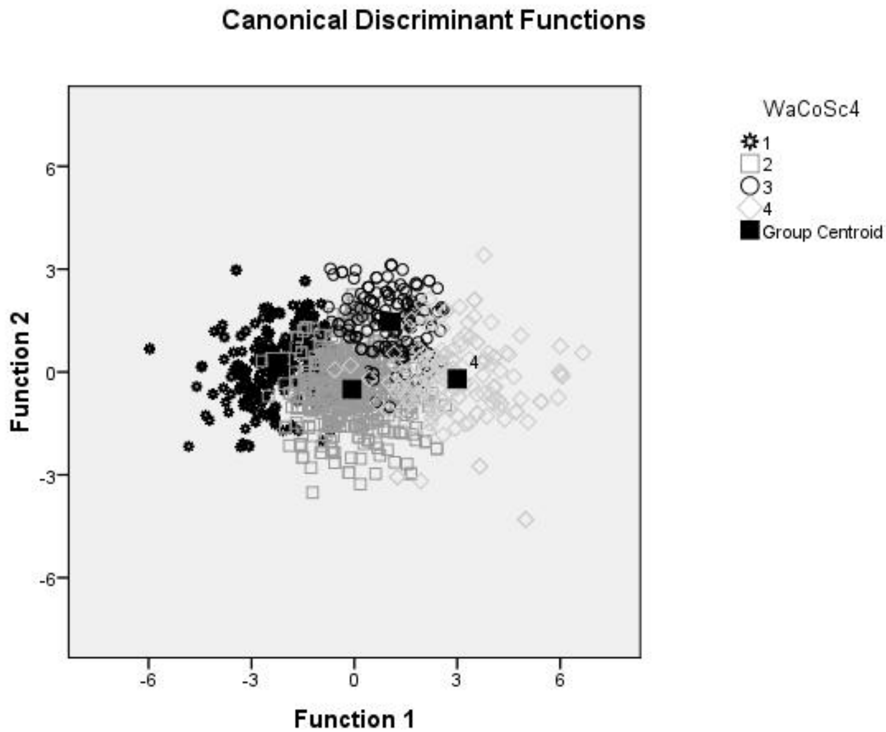


Figure 7. 5-Cluster DFA of Scaled Scores Using Ward's Method

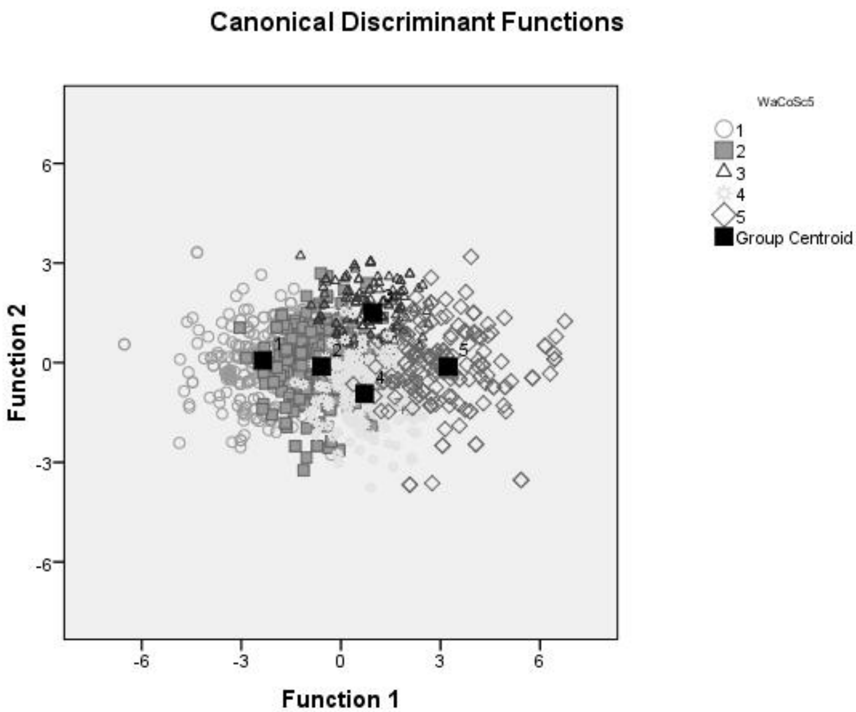


Figure 8. 6-Cluster DFA of Scaled Scores Using Ward's Method

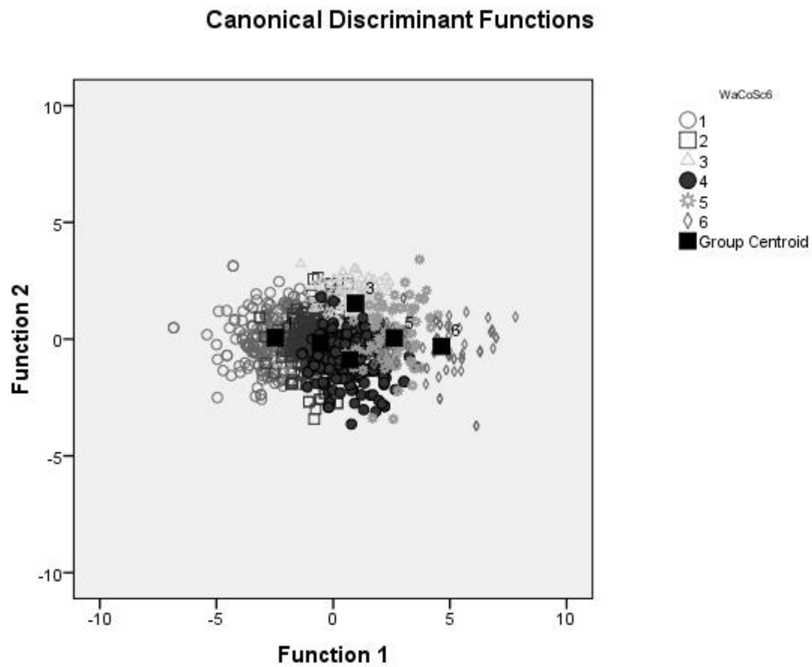
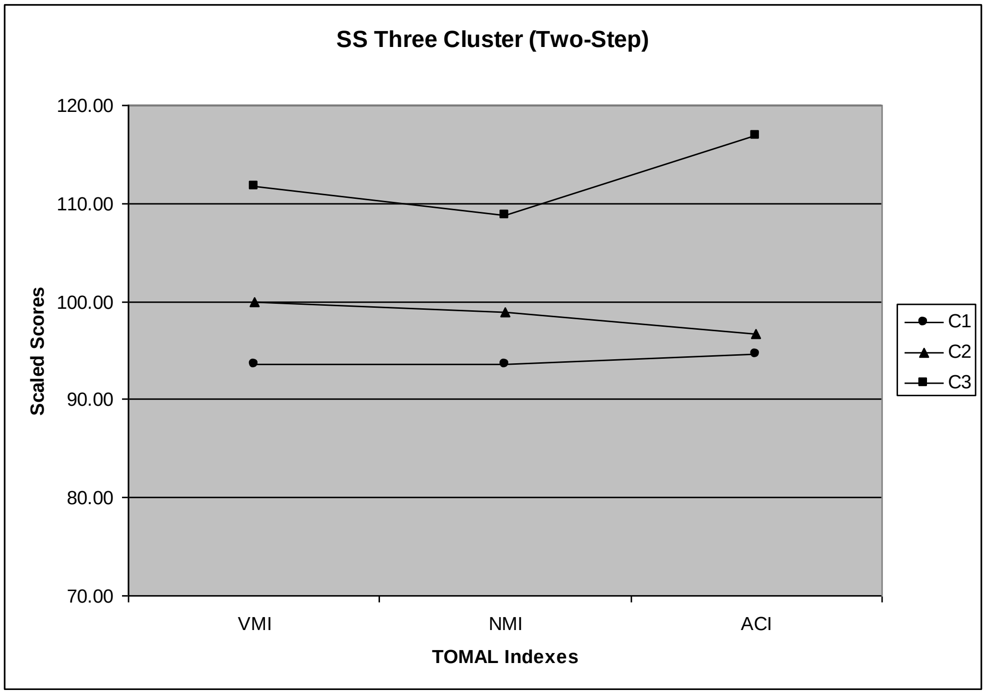


Figure 9. Cluster Profiles of TOMAL Indexes with a Three Cluster Solution: Raw Scores and Two-Step Method



Note. VMI = Verbal Memory Index, NMI = Nonverbal Memory Index, ACI = Attention/Concentration Index.

Note. C1 = Older Age Group, C2 = Middle Age Group, C3 = Younger Age Group

Figure 10. 3-Cluster DFA of Raw Scores using Two-Step Method

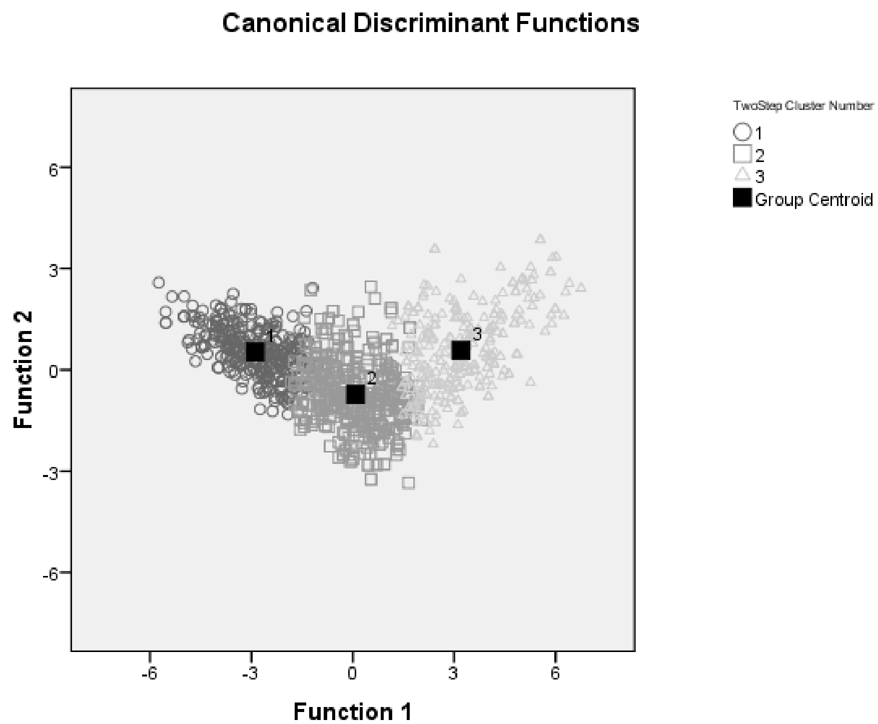
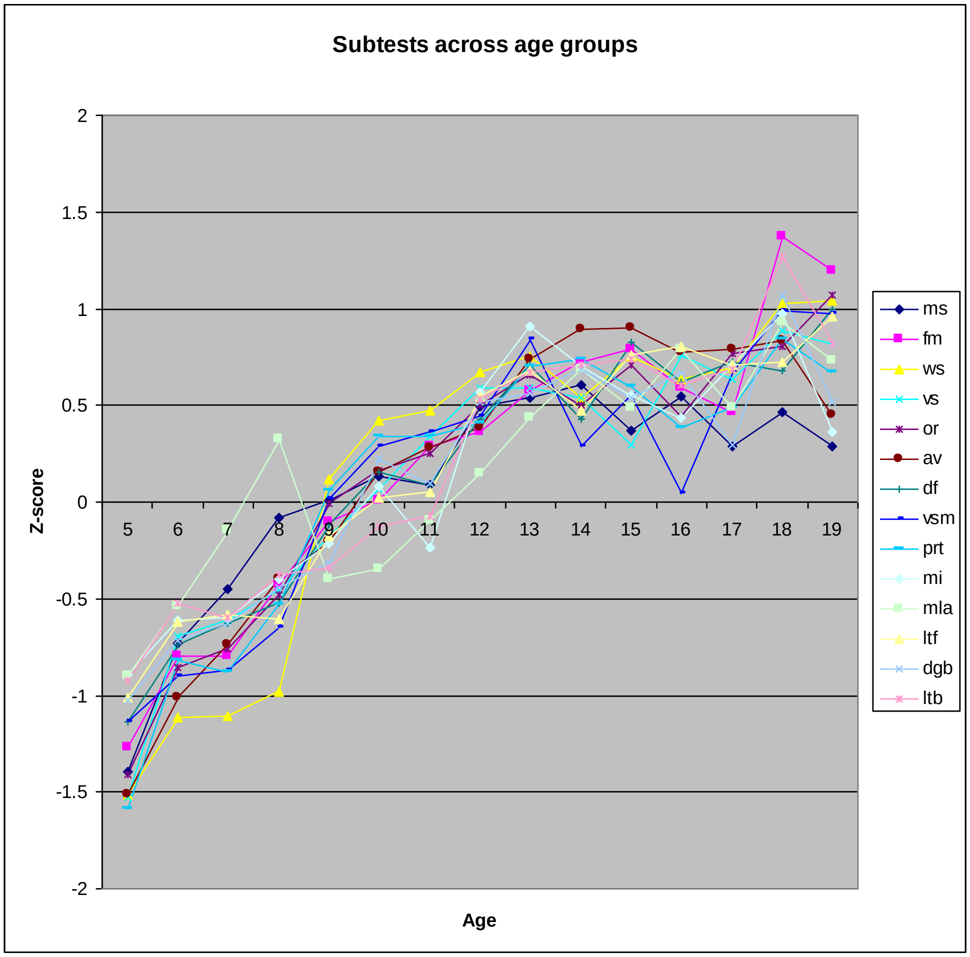
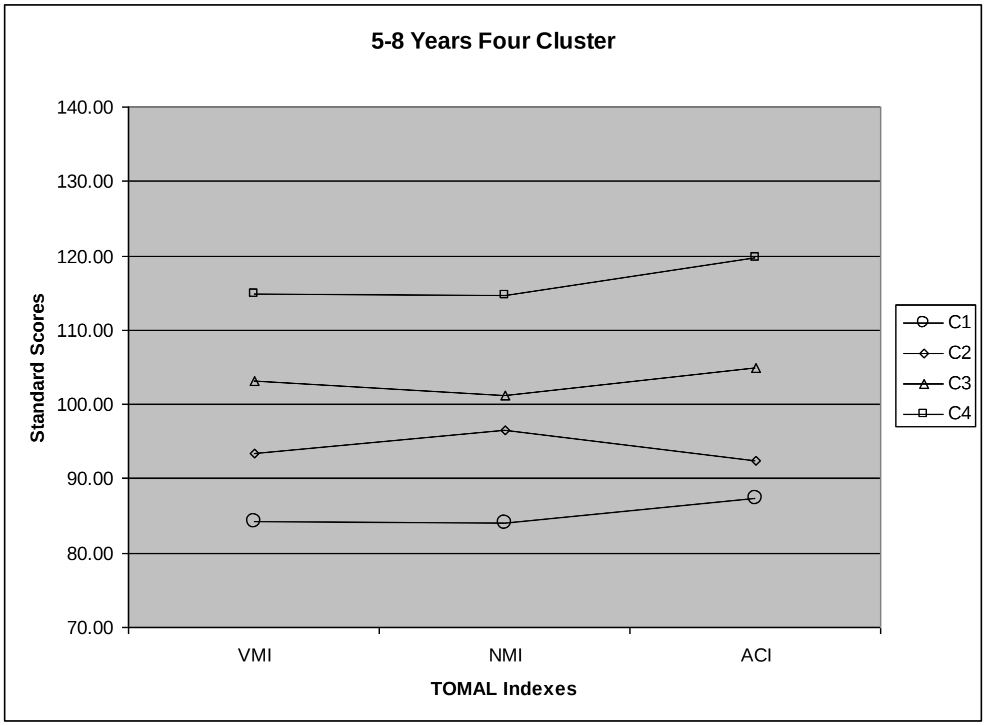


Figure 11. Subtest raw scores plotted across age groups.



Note. MFS = Memory for Stories, FM = Facial Memory, WSR = Word Selective Reminding, VSR = Visual Selective Reminding, OR = Object Recall, AVM = Abstract Visual Memory, DF = Digits Forward, VSM = Visual Sequential Memory, PR = Paired Recall, MLA = Memory for Location, MI = Manual Imitation, LF = Letters Forward, DB = Digits Backward, LB = Letters Backward.

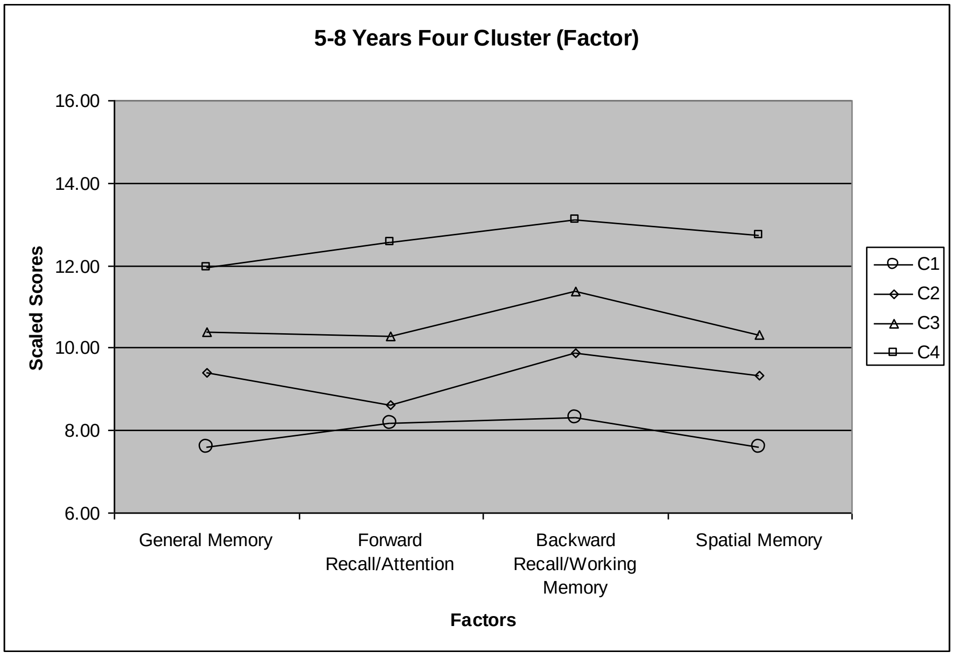
Figure 12. Cluster Profiles of TOMAL Indexes for Four Cluster Solution – 5-8 year group only: Scaled Scores and Ward's Method.



Note. VMI = Verbal Memory Index, NMI = Nonverbal Memory Index, ACI = Attention/Concentration Index.

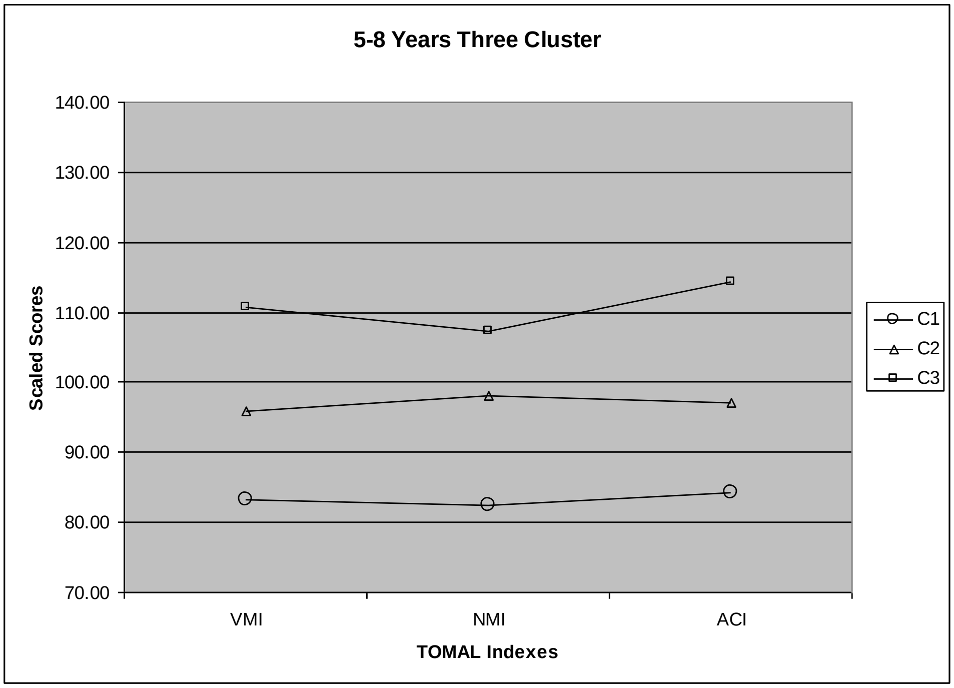
Note. C1 = High Average, C2 = Slightly Below Average, C3 = Slightly Above Average, C4 = Low Average

Figure 13. Cluster Profiles of Factor Scores for the Four Cluster Solution – 5-8 year group only: Scaled Scores and Ward's Method.



Note. C1 = High Average, C2 = Slightly Below Average, C3 = Slightly Above Average, C4 = Low Average

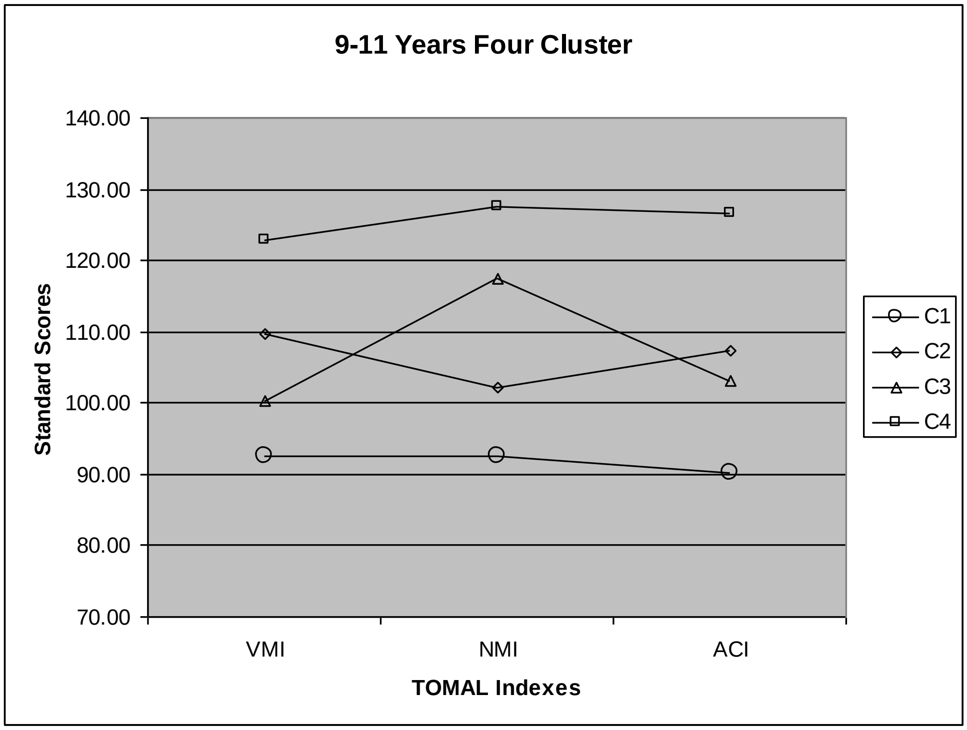
Figure 14. Cluster Profiles of TOMAL Indexes for Three Cluster Solution – 5-8 year group only: Scaled Scores and Ward's Method.



Note. VMI = Verbal Memory Index, NMI = Nonverbal Memory Index, ACI = Attention/Concentration Index.

Note. C1 = Low Average, C2 = Average, C3 = High Average

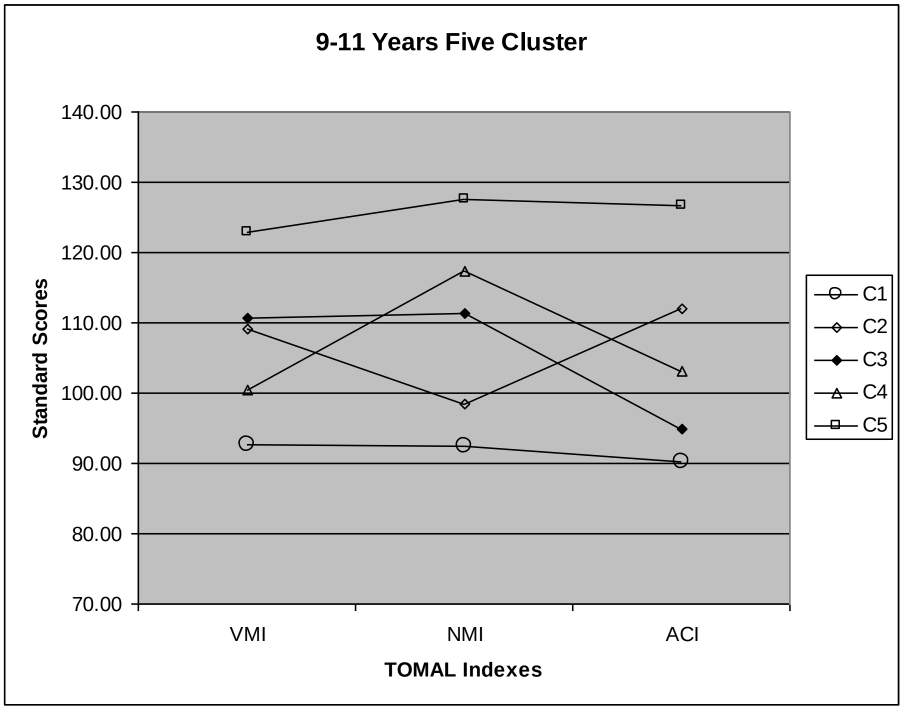
Figure 15. Cluster Profiles of TOMAL Indexes for the Four Cluster Solution – 9-11 years only: Scaled Scores and Ward's Method.



Note. VMI = Verbal Memory Index, NMI = Nonverbal Memory Index, ACI = Attention/Concentration Index.

Note. C1 = Low Average, C2 = Average, C3 = Nonverbal, C4 = High Average

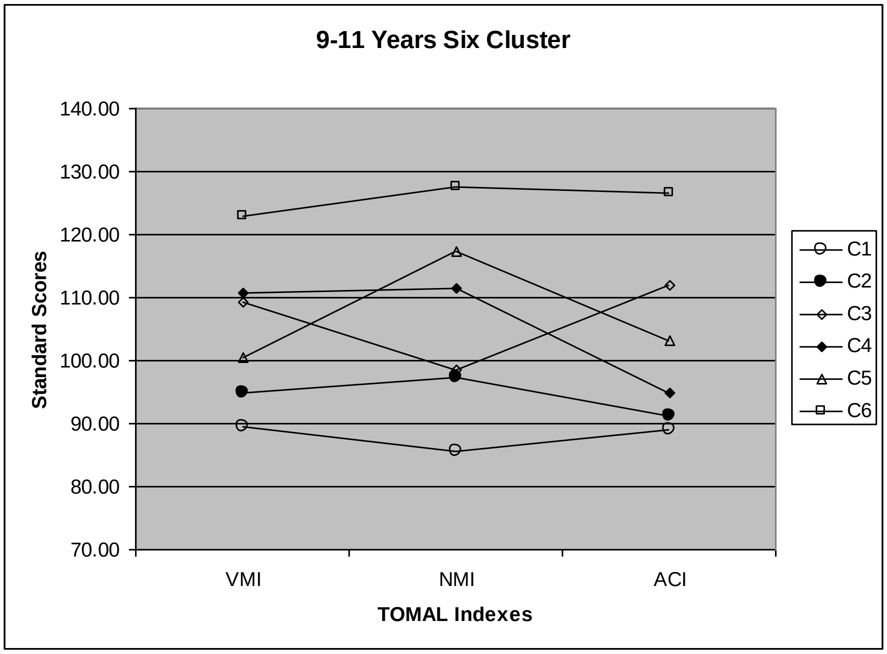
Figure 16. Cluster Profiles of TOMAL Indexes for the Five Cluster Solution – 9-11 years only: Scaled Scores and Ward's Method.



Note. VMI = Verbal Memory Index, NMI = Nonverbal Memory Index, ACI = Attention/Concentration Index.

Note. C1 = Low Average, C2 = Attention/Verbal, C3 = General Memory, C4 = Nonverbal, C5 = High Average

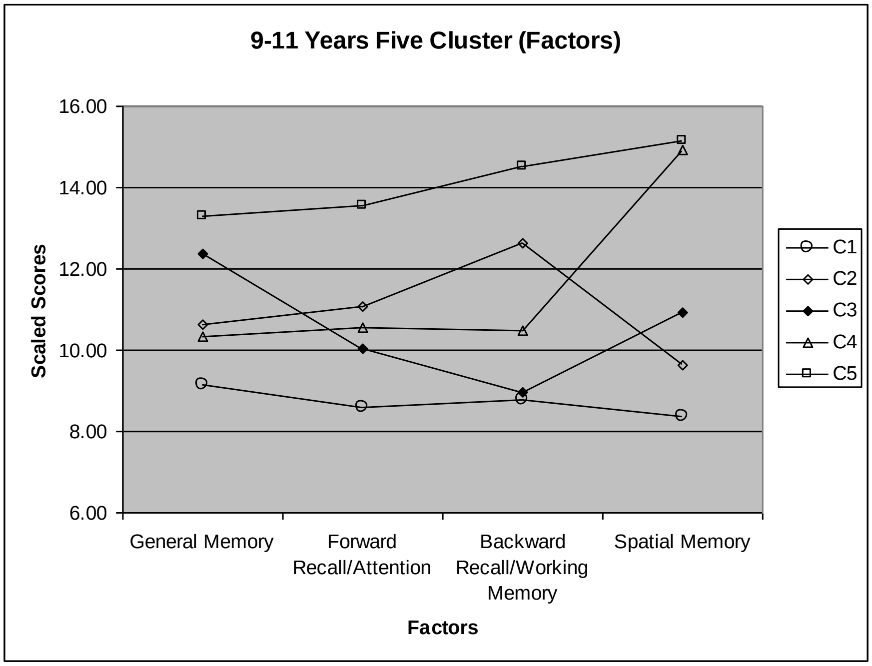
Figure 17. Cluster Profiles of TOMAL Indexes for the Six Cluster Solution – 9-11 years only: Scaled Scores and Ward's Method.



Note. VMI = Verbal Memory Index, NMI = Nonverbal Memory Index, ACI = Attention/Concentration Index.

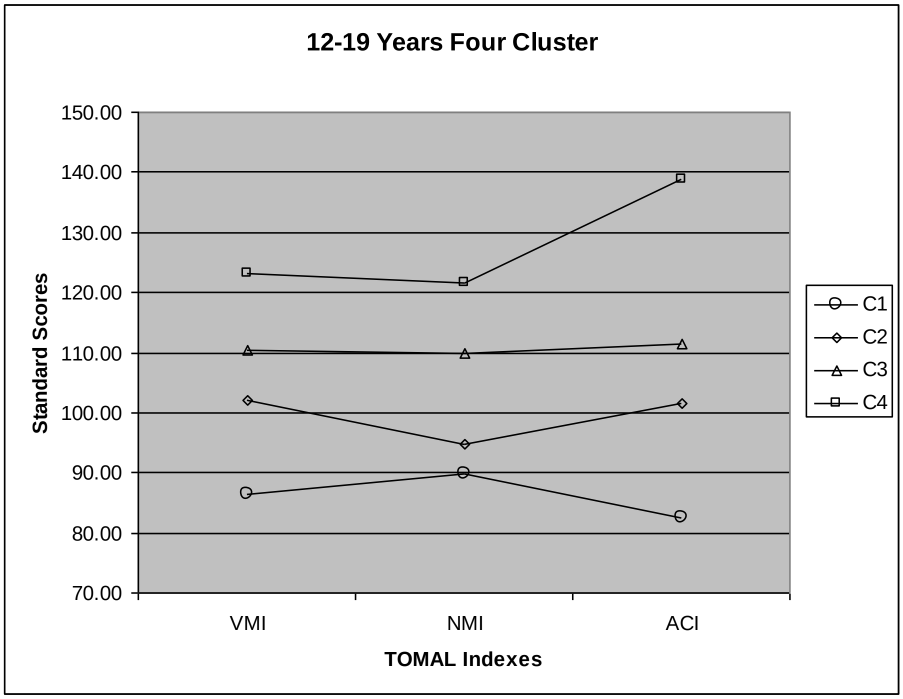
Note. C1 = Low Average, C2 = Average, C3 = Attention/Verbal, C4 = General Memory, C5 = Nonverbal, C6 = High Average

Figure 18. Cluster Profiles of Factor Scores for the Five Cluster Solution – 9-11 age group only: Scaled Scores and Ward's Method.



Note. C1 = Low Average, C2 = Attention/Verbal, C3 = General Memory, C4 = Nonverbal, C5 = High Average

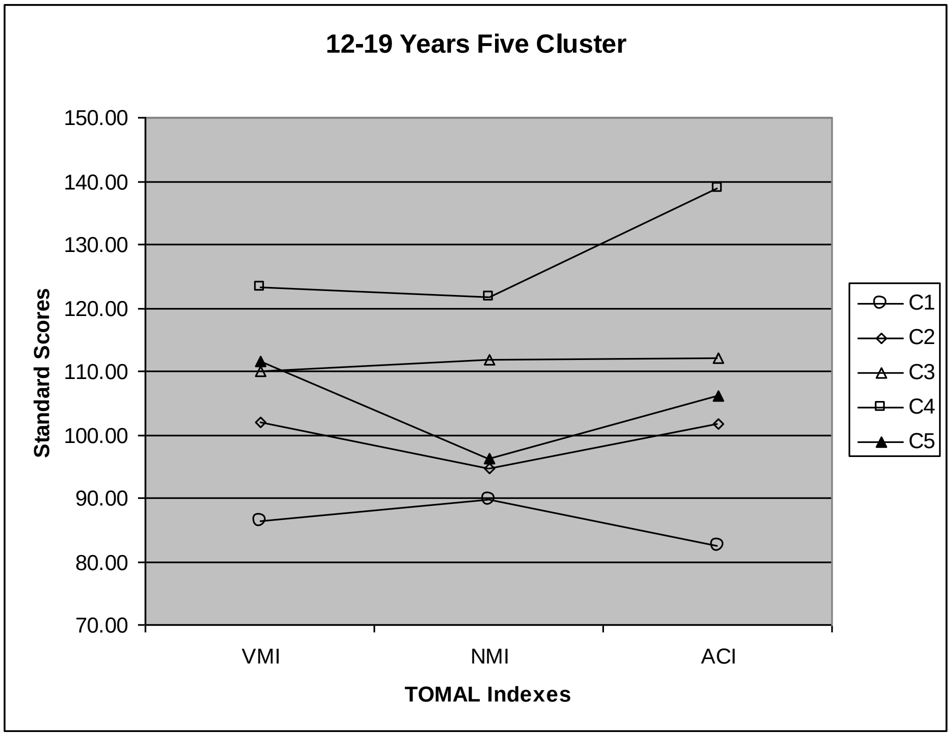
Figure 19. Cluster Profiles of TOMAL Indexes for the Four Cluster Solution – 12-19 years only: Scaled Scores and Ward's Method.



Note. VMI = Verbal Memory Index, NMI = Nonverbal Memory Index, ACI = Attention/Concentration Index.

Note. C1 = Low Average, C2 = Average, C3 = High Average, C4 = Advanced

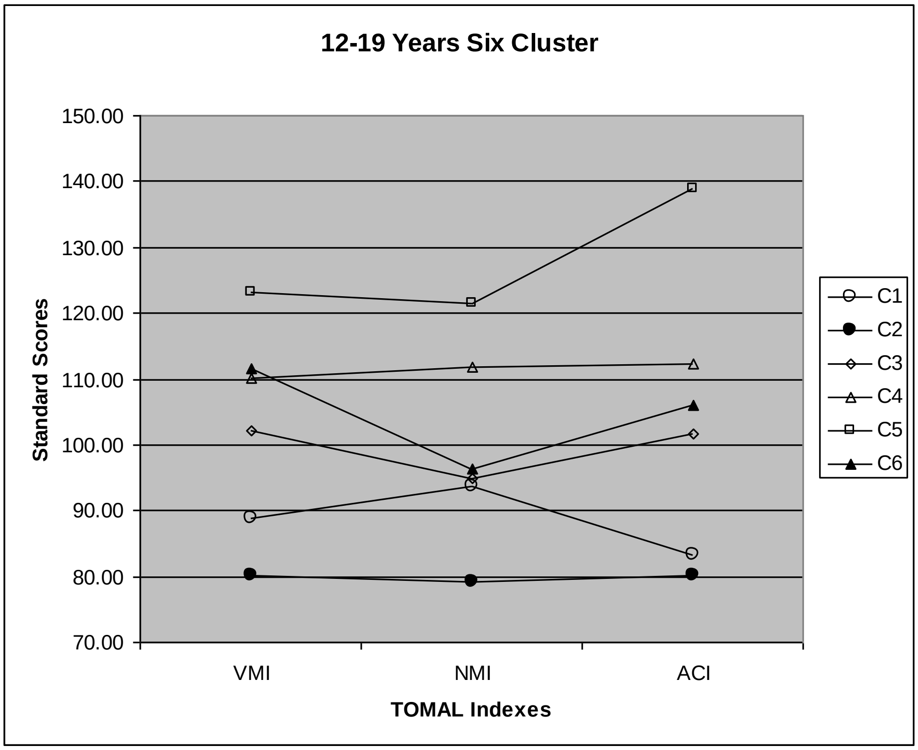
Figure 20. Cluster Profiles of TOMAL Indexes for the Five Cluster Solution – 12-19 years only: Scaled Scores and Ward's Method.



Note. VMI = Verbal Memory Index, NMI = Nonverbal Memory Index, ACI = Attention/Concentration Index.

Note. C1 = Low Average, C2 = Average, C3 = High Average, C4 = Advanced, C5 = Verbal

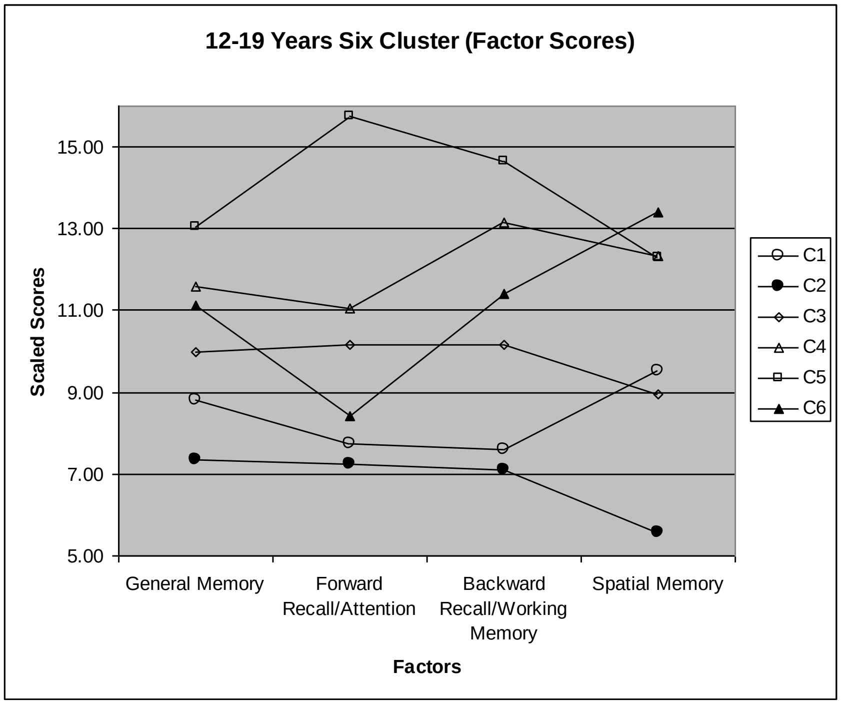
Figure 21. Cluster Profiles of TOMAL Indexes for the Six Cluster Solution – 12-19 Group Only: Scaled Scores and Ward's Method.



Note. VMI = Verbal Memory Index, NMI = Nonverbal Memory Index, ACI = Attention/Concentration Index.

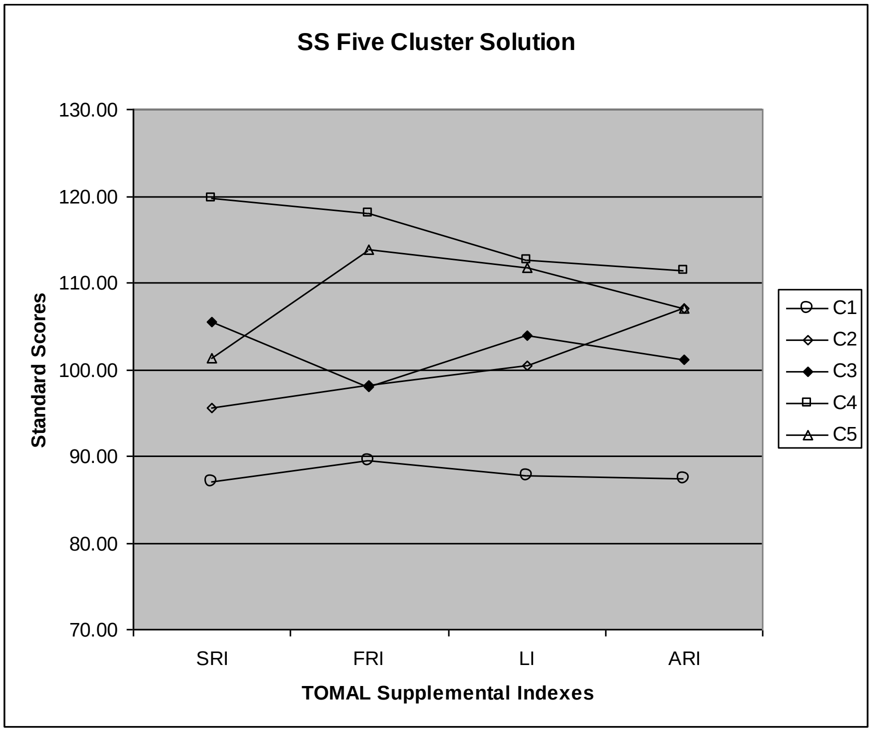
Note. C1 = Low Average, C2 = Low, C3 = Average, C4 = High Average, C5 = Advanced, C6 = Verbal

Figure 22. Cluster Profiles of Factor Scores for the Six Cluster Solution – 12-19 years only: Scaled Scores and Ward's Method.



Note. C1 = Low Average, C2 = Low, C3 = Average, C4 = High Average, C5 = Advanced, C6 = Verbal

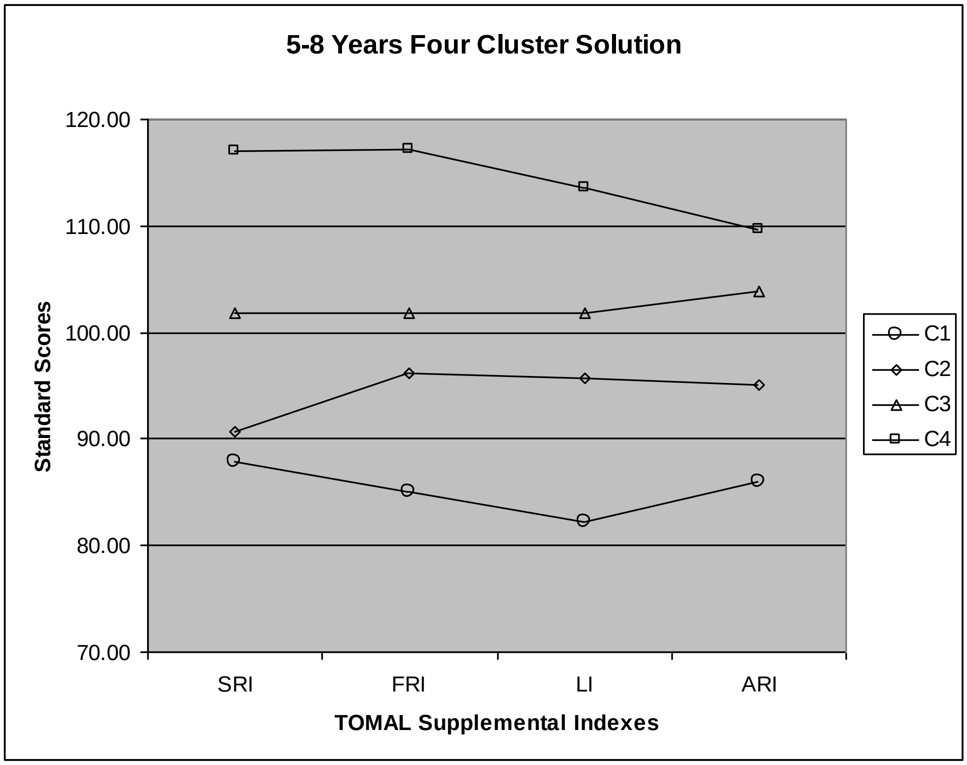
Figure 23. Cluster Profiles of Supplemental Index Scores for the TOMAL Standardization Sample: Five-Cluster Solution



Note. SRI = Sequential Recall Index, FRI = Free Recall Index, LI = Learning Index, ARI = Associative Recall Index.

Note. C1 = Low Average, C2 = Average, C3 = Verbal, C4 = High Average, C5 = Nonverbal

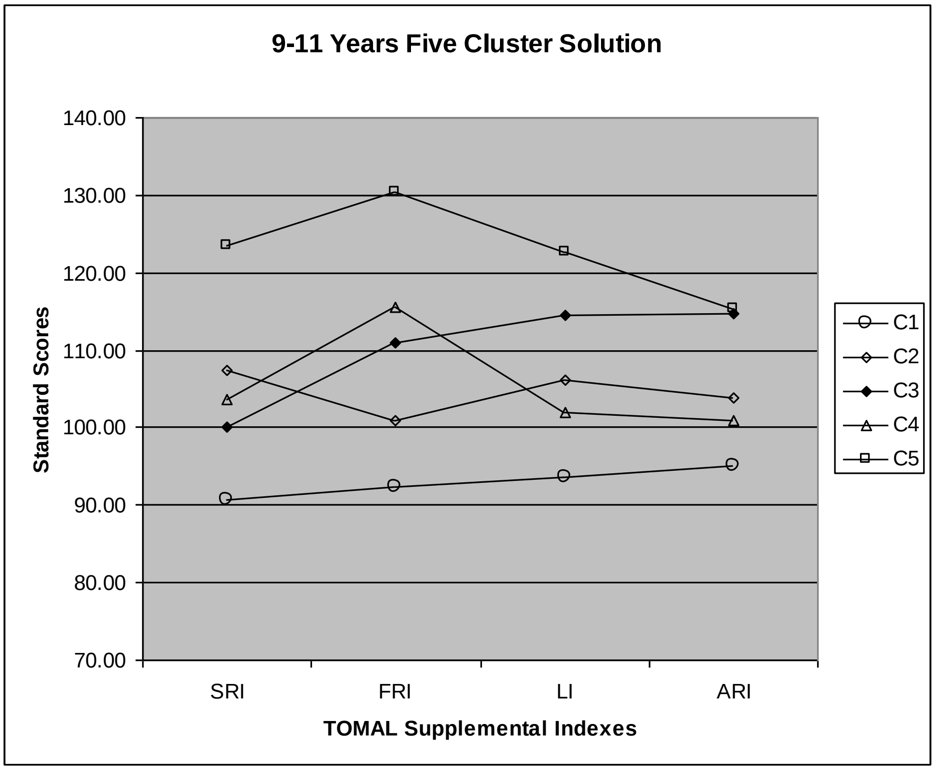
Figure 24. Cluster Profiles of Supplemental Index Scores for the 5-8 Years Group: Four-Cluster Solution



Note. SRI = Sequential Recall Index, FRI = Free Recall Index, LI = Learning Index, ARI = Associative Recall Index.

Note. C1 = High Average, C2 = Slightly Below Average, C3 = Slightly Above Average, C4 = Low Average

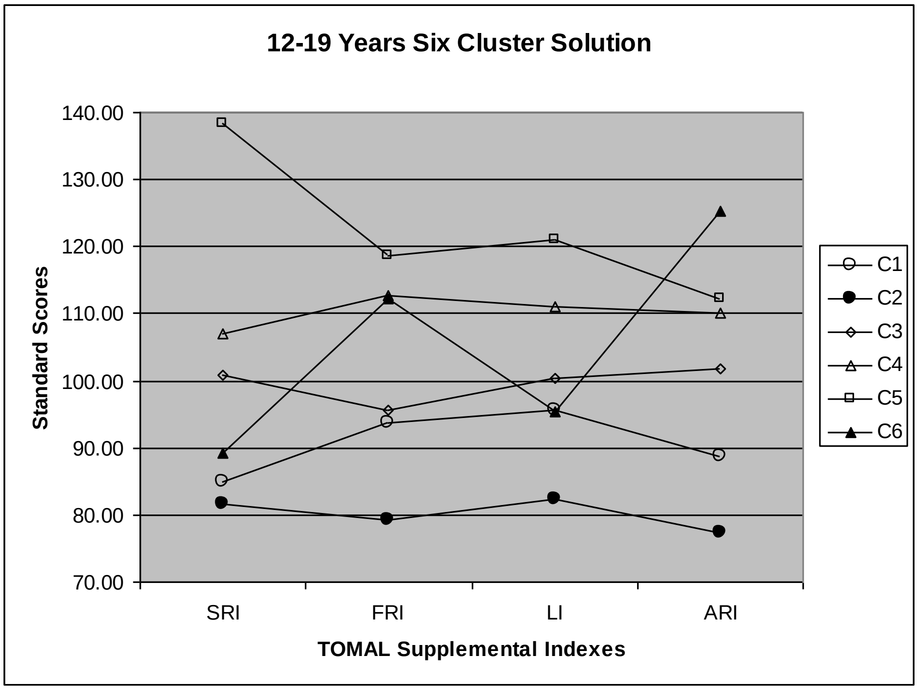
Figure 25. Cluster Profiles of Supplemental Index Scores for the 9-11 Years Group: Five-Cluster Solution



Note. SRI = Sequential Recall Index, FRI = Free Recall Index, LI = Learning Index, ARI = Associative Recall Index.

Note. C1 = Low Average, C2 = Attention/Verbal, C3 = General Memory, C4 = Nonverbal, C5 = High Average

Figure 26. Cluster Profiles of Supplemental Index Scores for the 12-19 Age Group: Six-Cluster Solution



Note. SRI = Sequential Recall Index, FRI = Free Recall Index, LI = Learning Index, ARI = Associative Recall Index.

Note. C1 = Low Average, C2 = Low, C3 = Average, C4 = High Average, C5 = Advanced, C6 = Verbal

VITA

Graduate College
University of Nevada, Las Vegas

Nicholas Shizuo Thaler

Degrees:

Bachelor of Science, Psychology, 2004
University of California, San Diego

Master of Arts, Clinical Psychology, 2006
California State University, Northridge

Publications:

Thaler, N. S., Kazemi, E., & Wood, J. J. (in press). Measuring anxiety in youth with learning disabilities: Reliability and validity of the Multidimensional Anxiety Scale for Children (MASC). *Child Psychiatry and Human Development.*

Thaler, N. S., Allen, D. N., Park, B. S., McMurray, J. C., & Mayfield, J. (in press). Attention processing abnormalities in children with TBI and ADHD: Differential impairment of component processes. *Journal of Clinical and Experimental Neuropsychology.*

Thaler, N. S., Allen, D. N., McMurray, J. C., & Mayfield, J. (2010). Sensitivity of the Test of Memory and Learning to attention and memory deficits in children with ADHD. *The Clinical Neuropsychologist, 24,* 246-264.

Allen, D. N., Thaler, N. S., Donohue, B., & Mayfield, J. (in press). WISC-IV profiles in children with traumatic brain injury: Similarities and differences to the WISC-III. *Psychological Assessment.*

Thaler, N. S., Kazemi, E., & Huschler, C. (2009). Developing a rubric to assess student learning outcomes using a class assignment. *Teaching of Psychology, 36*(2), 113-116.

Thesis Title: Cluster Analysis of the TOMAL Standardization Sample

Thesis Examination Committee:

Chairperson, Daniel N. Allen, PhD
Committee Member: Brad Donohue, PhD
Committee Member: Murray G. Millar, PhD
Graduate Faculty Representative: Chad Cross, PhD