CONSTRUCTION AND VALIDATION OF AN ECOLOGICAL MEASURE OF WORKING MEMORY

A Dissertation Submitted to the Temple University Graduate Board

---

In Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

---

By Gina A. Forchelli

May 2015

Examining Committee Members:
Catherine Fiorello, Advisory Chair, Psychological, Organizational, & Leadership Studies
Joseph Ducette, Psychological, Organizational, & Leadership Studies
Joseph Boyle, Psychological, Organizational, & Leadership Studies
Ken Thurman, Psychological, Organizational, & Leadership Studies
Frank Farley, Psychological, Organizational, & Leadership Studies

ABSTRACT

Working memory (WM) has been closely linked to learning and achievement in children (Gathercole et al., 2004). The Forchelli Following Directions Task (FFDT) is a 15-item group-administered screener designed to assess working memory ability in school-aged children. The FFDT was developed to address the need for early identification of children with working memory difficulty. It specifically focuses on the need for easily administered and ecologically valid assessment.

The FFDT was developed based on tasks cited in research to assess WM. The measure was developed across three iterations after receiving continual review from research experts in working memory and a group of three elementary school teachers. It also was piloted by three elementary school children to assess group-administration considerations. Participants in the validation study were 70 elementary school students 5 to 10 years of age spanning kindergarten to third grade were recruited from schools in the greater Philadelphia area. Participants were administered the group-administered working memory screener and completed individually administered measures of working memory, the WISC-IV Digit Span and Spatial Span, for comparison. Parents and teachers also completed behavior rating scales (i.e., BRIEF) measuring working memory.

The FFDT demonstrated a sufficient Alpha's coefficient, indicating internal consistency. Significant Pearson correlations were found between existing measures of WM and the FFDT, indicating that the FFDT measures WM ability to a similar extent. The FFDT demonstrated good sensitivity to age and grade, as well. Further, the results of a ROC analysis comparing the identification of WM difficulty on the FFDT to existing measures of WM demonstrated a low to moderate effect.

Overall, results indicate that the FFDT exhibited good reliability and validity. The anecdotal support of elementary school teachers and time efficiency of the task compared to existing WM measures also suggests good ecological validity. This study also demonstrated the utility of the FDDT in populations within a Response to Intervention (RtI) framework. Further research will be challenged to investigate the FFDT further scrutinize the construct validity and demonstrate significance in a larger, more representative sample of students.

ACKNOWLEDGEMENTS

I am sincerely grateful to the many individuals that have supported and guided me through this journey; their love and encouragement has allowed me successful completion of my dissertation project and my doctoral program. First, I want to extend a big thank you to my advisor, mentor, and dissertation chair, Dr. Catherine Fiorello. It has been an honor to learn and work with you these past four years and I would not have accomplished these successes without your assiduous attention to detail, guidance, and support.

Thank you to the rest of my dissertation committee, Dr. Joseph Boyle, Dr. Joseph DuCette, Dr. Frank Farley, and Dr. Ken Thurman. Your valuable insights and feedback have challenged me and allowed my work to evolve into a product I am proud of. I also extend thanks to the rest of the school psychology faculty, including Drs. Jean Boyer and Merideth Weber for your mentorship and support.

I would also like to extend my thanks to all the students, parents, teachers, and school administration who participated in this study. Their contributions were paramount to this achievement.

Most importantly, I want to thank the positive influences and supports in my personal life. I would not have been able to push through the many hardships throughout this endeavor without the support of my brothers, Vincent Forchelli and Christopher Forchelli and, best friend, Ashley Benjamin. I valued your ability to help me laugh throughout difficult situations and share in the celebration of my successes. I extend the most gratitude to my biggest fans, my parents, Elizabeth and John Forchelli. From an

early age, you have instilled in me the importance and value of knowledge. It is this inner drive that has afforded me the courage to attempt, as well as the grit and stamina to, successfully complete this achievement.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE

INTRODUCTION

Context

Working memory (WM) is defined as the ability for one to temporarily store and manipulate or transform presented auditory and visual information (Baddeley, 2000). It is a useful scientific construct that is important in a wide range of cognitive functions and behaviors (Conway, Kane, Bunting, Hambrick, Wilhelm, & Engle, 2005). Additionally, assessment of WM in clinical and educational settings is essential when assessing for learning difficulties (Rogers, Hwang, Toplak, Weiss, & Tannock, 2011).

WM develops gradually through childhood (Brocki & Bohlin, 2004; Gathercole et al., 2004; Wu et al., 2011). It has been demonstrated in children as young as four years old (Alloway et al., 2004; Gathercole et al., 2004) and becomes more sophisticated and efficient through adolescence (Cocklin et al., 2007; Lambek and Shevlin, 2011). It also has been closely linked to the development of language (Adams & Gathercole, 1995; Baddeley, 1996; Botting & Conti-Romseden, 2001; Montgomery et al., 2009), attention (Alloway, 2011: Martinessen et al., 2005; Raiker et al., 2012), and executive functions (Barkley, 1997, 2001; Brocki & Bohlin, 2004, Latzman & Markon, 2009; Miyake et al., 2000; Wu et al., 2011), all of which are integral to learning new information.

Recent research has demonstrated a strong link between WM and academic achievement; that is, children with poorer WM tend to perform poorly on academic tasks (Gathercole, Pickering, Knight, & Stegmann, 2004). This has been linked to higher-level cognitive tasks, such as reading comprehension, language comprehension, and mathematics (Conway, Kane, Bunting, Hambrick, Wilhelm, & Engle, 2005) and demonstrated in elementary school children through adolescence (Gathercole & Pickering, 2004; Rogers et al., 2011). In addition, this phenomenon is not isolated to children with learning disabilities (LD); it affects children with and without a diagnosis of LD (Pickering, 2006).

This suggests that WM measurement is important in identifying and supporting children within their academic setting. WM skills are typically assessed through “span tasks,” requiring the individual to recall and transform an increasing number of sequenced items (Gathercole & Alloway, 2008). These span tasks are among the most utilized assessments in psychology and are widely considered reliable and valid (Conway et al., 2005).

Recently, it has been argued that a task requiring a child to repeat a span of non-meaningful items, such as letters or numbers, is not an ecologically valid measure of the WM skills necessary within the classroom (Levin, Thurman, & Kiepert, 2010). According to Levin, Thurman, and Kiepert, for a task to provide meaningful information, it should appear similar to tasks performed in the everyday environment. Similarly, Spooner and Pachana (2006) have discussed the importance of the task to predict the performance of similar tasks in real-world environments.

Levin, Thurman, and Kiepert (2010) discuss two ways of assessing WM in a more ecologically valid way; namely, they discuss behavioral observations in the natural environment and classroom-based performance measures. Due to the issues surrounding behavior rating scales, including lack of sensitivity in general behavior (Kolko & Kazdin, 1998) and specific WM behavior (Norman & Tannock, 2012), classroom-based performance measures appear to be a better approach. These performance tasks attempt to mirror classroom activities that utilize WM. For example, tasks requiring children to following multi-step directions similar to those heard in the classroom are argued to tax WM resources. In studies utilizing a following directions task, they were found to have a relationship to standard WM tasks, as well as a relationship to academic performance (Gathercole et al., 2007; St. Clair-Thompson et al., 2010). However, these measures have only been utilized in few research studies, as outcome measures of an intervention, rather than as clinical measures.

Further, a measure is needed to respond to changes in how identification and intervention services are provided. Wass and colleagues (2012) suggest that early WM intervention is related to widespread transfer of skills. There also has been a move in the American educational system to earlier identification and continuous progress monitoring of academic difficulties (Brown-Chidsey and Steege, 2005), entitled Response to Intervention (RtI). In this model, students are screened earlier for academic difficulties and provided interventions to remediate difficulty. However, there are limited existing screeners utilized to identify and monitor cognitive (i.e., WM) growth and development.

Statement of Purpose

Working memory (WM) is a construct that is closely related to learning and achievement in children (Gathercole et al., 2004; Pickering, 2006). Children with poor WM display more academic difficulties which appear to worsen over time without intervention (Alloway et al., 2009). However, with intervention, WM ability can improve, and there is some indication that these improvements transfer to academic achievement (Holmes et al., 2009). Therefore, identification of WM difficulty is important, with earlier identification providing better remediation effects (Wass, 2012).

Ecologically valid assessment provides meaningful information about functioning (Levin, Thurman, & Kiepert, 2010); however, there are limited measures that utilize ecologically valid methods to assess WM. These few measures have been criticized for lack of concurrent validity; that is, measures are not demonstrating strong correlations with related measures of working memory (Kolko & Kazdin, 1998; Norman & Tannock, 2012). Furthermore, there is a need to develop measures that increase feasibility and "goodness of fit" in the American educational system. Therefore, it is imperative to investigate alternative measures of WM that fit these criteria.

The purpose of this study is to develop and examine the validity of a group-administered ecologically valid measure of WM in school-aged children. It will allow for earlier and more efficient identification of students with difficulties and can be generalized to American students. This study defined the construct of WM through an extensive literature review and attempted to validate the following directions task through measurement of construct validity.

Research Questions

1. Is an ecologically-valid working memory measure (i.e., following directions task) related to existing, validated measures of working memory?

The developed task was administered to school-aged children and compared to standard measures of WM, including a span task and a behavior rating scale. The WISC-IV Digit Span subtest was utilized as the span task and the teacher and parent rating forms of the BRIEF will be utilized as the behavioral measure of WM. Past research has provided some evidence that a task requiring children to following instructions is indicative of WM ability (Gathercole et al., 2008; St. Clair-Thompson et al., 2010). It is hypothesized that students will perform similarly on the developed following directions task to the more standard measure of WM.

2. Is an ecologically-valid working memory measure (i.e., following directions task) more strongly related to existing measures of WM than behavioral rating scales?

Behavioral screeners of internal states have been criticized in research (Kolko & Kazdin, 1998; Norman & Tennock, 2012). This suggests that a performance-based measure is more effective at indicating difficulty in cognitive states, such as WM. It is hypothesized that the following directions task will have a stronger relationship with the standard WISC-IV Digit Span subtest than BRIEF observant-rater forms completed by students' teachers and parents.

Definition of Terms

Short-term Memory (STM) is the ability to hold and perform rote recall after a short delay (Miller, 1956). In Baddeley's (2000) WM model, the phonological loop is considered the verbal STM and the visual-spatial sketchpad is considered the catch-all for visual and non-verbal STM. Studies interchange these terms (i.e., phonological loop and visual-spatial sketchpad) with STM; however, verbal and visual STM will be used in this review.

Working Memory (WM) is the cognitive system that temporarily stores and manipulates auditory and visual information (Baddeley, 2000). This encompasses the use of STM systems (i.e., phonological loop and visual-spatial sketchpad), as well as more higher-order systems, namely the central executive and the episodic buffer, to rehearse stored information, connect it to other relevant information in LTM, and transform it into a new product.

Central Executive (CE) is considered the attention subsystem of the WM system (Baddeley, 2000). It allocates attention to the varying parts of the WM system. Many studies utilize this term to describe difficult WM tasks; however, it is misrepresentative to identify an entire system by only one part of the system. In this review, it is assumed that a task that is measuring "CE" is actually measuring the WM system and, therefore, will be entitled a WM task.

Episodic Buffer (EB) is considered the communication subsystem, that links information between the STM and LTM in order to decide what information is meaningful and takes priority (Baddeley, 2000). It binds and holds the information held in STM for use in a task or further processes into LTM.

Long-term Memory (LTM) is considered the long-term storage that represents an individual's knowledge of the world (Hunt & Ellis, 2004). It represents the deeper and more elaborate processing that leads to better learning of information.

STM Tasks are considered tasks that researchers believe tap more into the slave systems of the WM system, including the phonological loop and visual-spatial sketchpad. A few popular tasks include the following:

1. *Forward Span Tasks* require participants to listen to a sequence of letters, numbers or objects and are asked to repeat the sequence in forward sequence. This can also be presented visually utilizing a board or grid where an individual follows and repeats a visual pattern in forward sequence.

WM Tasks are considered tasks that researchers believe tap into the auditory system of WM. A few popular tasks include the following:

1. *Backward Span Tasks* require participants to listen to sequence of letters, numbers or objects and are asked to repeat the sequence in the reverse or some otherwise ordered sequence. This can also be presented visually utilizing a board or grid where an individual follows and repeats a visual pattern in a similar manner.
2. *Sentence Recall Task* requires a participant to listen to a group of sentences. After each sentence, the participant is required to state if the sentence if true or false. Then, at the end of each group of sentences, they are required to remember the last word of each sentence stated.

CHAPTER TWO

REVIEW OF THE LITERATURE

Working Memory

Working memory (WM) is defined as the cognitive system that temporarily stores and manipulates auditory and visual information (Baddeley, 1986, 2000; Baddeley & Hitch, 1974). Unlike short-term memory (STM), WM is designed to hold and connect pieces of information to execute higher-order cognitive processes. It allows individuals the opportunity to plan and carry out complex instructions and problem solve. It is closely linked to one's ability to comprehend and articulate language (Boudreau & Costanza-Smith, 2011) and relates to one's ability to understand complex situations (Sweller, van Merrienboer, & Paas, 1998).

WM was first discussed by Miller, Galanter, and Pribram (1960) in an attempt to explain how internal concepts work with external stimuli to formulate new information. Miller (1956) determined that a STM store could hold approximately seven units of information for rote recall. Miller, Galanter, and Pribram (1960) argued that this limited-capacity store could not feasibly hold all of the rules or sub-goals to complete more complex tasks. For example, to solve a mathematical equation in one's head, an individual must simultaneously shift between the concepts surrounding numeracy, and the rules of the equation, as well as refreshing or transforming the equation as conclusions are reached. The STM store could not sustain all this information at once. Therefore, a different system, entitled WM, must be present that juggles these various processes necessary to complete the task.

There are two main approaches by which theorists have conceptualized WM, either as a unitary-store or multi-store system. Both theoretical viewpoints agree that information is processed in modality-specific ways; that is, verbal and visual information. However, they differ in their argument for where and how the modal information is processed. In the unitary-store models, all information in the short-term memory (STM) and long-term memory (LTM) are processed within the same anatomical system. Conversely, in the multi-store models, information is processed by differing systems or pathways.

Unitary-store theorists have postulated that WM resides within one larger memory system. The embedded processes theory by Cowan (1988; 1999; 2001) hypothesized that all information, no matter the modality, was processed in the LTM system. In this approach, information enters the memory system by way of focused attention; attended information is processed, unattended information is not. Once focused upon, the information activates existing knowledge representations within the LTM system. These active units are based upon sensory and semantic characteristics of the incoming information. For example, stating “red robin” may activate memory related to the visual experience of the bird and a verbal experience of the word. These different modalities of the information are activated in the same place. Then, the chosen activated parts of the LTM transform the incoming information into meaningful units, also known as binding; this process is considered the WM of the system. Recent unitary-store theorists have built upon Cowan’s (1988) model to include hierarchical levels of processing within the focused attention to accommodate a larger processing capacity (Oberauer, 2009); however, they all assume a unitary unit.

One of the most widely used theories is the multi-component WM theory (Baddeley, 2000). Baddeley divides the WM system into four components: the phonological loop, the visual-spatial sketchpad, the central executive, and the episodic buffer. When information is first presented to the WM system, the central executive (CE) determines what information will be processed. Baddeley (2012) describes the CE as the "homunculus" (p.13) of the WM system. This is the attentional capacity of the WM system; it includes focused, divided, and switching attention (Baddeley, 1996). Information that is attended to is temporarily stored by stimuli specific “slave” systems; the phonological loop and the visual spatial sketchpad. The phonological loop stores all language-based information, including both vocal and sub-vocal information. The visual-spatial sketchpad processes all incoming visual and spatial information.

The CE determines how to attend to information through continued communication with the fourth component, the episodic buffer. The buffer communicates with the perceptual and long-term memory (LTM) systems in order to decide what information is meaningful and takes priority (Baddeley, 2000). It, then, integrates, or binds, and holds the incoming auditory and visual information for use in a task or further processing into LTM.

Studies have attempted to demonstrate support for this structure of WM. Applied behavioral and neuropsychological evidence, by way of administering memory span tasks, have demonstrated strong evidence that memory is composed of multiple systems. Neurological studies that, by way of neuroimaging data, directly measure brain activity during memory tasks have shown mixed results in the structure of memory.

One way to behaviorally investigate memory components is through dual-task experiments, where an individual is required to complete two tasks that both compete for memory capacity. In a classic irrelevant speech experiment by Logie and Baddeley (1987), individuals were asked to count the number of times a square appeared on a computer screen while performing a competing task. If the competing task was auditory (i.e., saying the word "the"), the ability to monitor accurate counting was significantly impaired; however, if the competing task was visual/motor (i.e., finger tapping), performance was not significantly impacted. Similarly, in a separate experiment by Baddeley, Grant, Wight, and Thomson (1975), participants' ability to recall visually-relevant sentences (i.e., using "up", "down", "left", "right") during a visual-spatial tracking task was significantly impaired; however, if sentences lacked visually- or spatially- based information, performance on the tracking task was not impaired. If information were processed in a unitary system, all interference should equally disrupt recall; however, it appears that interference has modality-specific tendencies. These irrelevant speech experiments demonstrate that stimuli-specific interference disrupts the phonological loop's ability to hold and monitor auditory information, indicating independent slave systems. Recent replication studies continue to demonstrate capacity limitations to encoding visual information when auditory information was present (Dell'Acqua & Jolicoeur, 2000; Morey & Cowan, 2004; Stevanovski & Jolicoeur, 2007),

Individual difference studies also support a multi-component WM. In a recent study by Dagan and colleagues (2012), 348 college students were administered visual and verbal WM and an intelligence battery. They reported that visual and verbal WM differentially correlated with the students' cognitive abilities; that is, visual WM demonstrated strong correlations with measures of fluid reasoning while verbal WM demonstrated strong correlations with measures of crystallized knowledge. Individuals appear to demonstrate differential WM abilities, dependent on modality. This supports a domain-specific WM system.

Neuropsychological studies looked into the WM capacity of memory impaired participants in order to investigate memory functioning. Vallar and Baddeley (1984) suggested that the STM and WM are independent systems. In this study, a woman with STM loss had less difficulty recalling more complex words than single consonants; she had the ability to recall words of increasing length at a higher rate than consonants. Although her ability to process small units of language-based information was impaired, she still was able to process larger, semantically relevant units of information. Baddeley (2000) argues that these data support the existence of the episodic buffer, which integrates information across multiple avenues.

Similarly, studies have also demonstrated that memory impaired individuals with rote recall difficulties do not necessarily have difficulties with other areas of memory or learning (Shallice & Warrington, 1970; Warrington & Shallice, 1969). More recent research with a developmental amnesiac patient also demonstrate that the capacity to perform well on complex WM tasks is not determined solely by STM abilities (Baddeley, Allen, & Hitch, 2011; Baddeley, Allen, & Vargha-Khadem, 2010).

Data measuring brain activity, otherwise known as neuroimaging data, allow researchers to locate potential areas within the brain that are utilized during WM tasks. These data demonstrate mixed results for the existence of independent areas in WM. Smith and Jonides (1997; 1999) demonstrated that the left supramarginal gyrus and left frontal lobe (i.e., Broca's area) are activated when processing language-based information, indicative of the phonological loop. There appears to be more division of labor in the visuospatial sketchpad; object and spatial information appear to be differentially processed in the ventral and dorsal visual streams of the brain, respectively. In these same tasks, when the complexity of the task increases, so does the activation of the dorsolateral prefrontal cortex (DLPFC), providing some evidence of where the CE is located.

However, recent neuroimaging data have also suggested single pathway or unit activation when performing WM tasks. More specifically, some studies show that STM tasks often activate areas of the brain involved in the LTM, suggesting that they are all part of the similar network (Jonides et al., 2008). Research has also suggested that a specific unitary region, the left Intraparietal Sulcus (IPS) is the main area activated during multi-modal WM tasks (Cowan et al., 2011); however, the study was somewhat weak. The lower sample size of sixteen participants from ages 18 to 24 years is not representative enough to draw definitive conclusions. Furthermore, in a follow-up study cited in the same article, there was also right IPS activation during a similar task. It does not appear that there is enough information to assume that one unitary area is responsible for WM.

In general, neuro-imaging data should be interpreted with caution. Many researchers have critiqued the use of neuro-imaging data to draw concrete conclusions about behavior and functioning (Beck, 2010; Dobbs, 2005). It is unclear that the neuro-imaging instruments truly measure functioning; that is, the scientific community cannot prove that a correlation between high activity in a specific area of the brain, measured by oxygen levels or blood-flow in the brain, and a specific behavior indicates that the brain activity causes the behavior observed. In WM research, the activation of an area of the brain during a recall task does not equate the necessity of that area to the task. For example, when asked to remember and repeat a word, an individual would need to activate areas responsible for phonological, articulatory, and lexical aspects of that word that may not be necessary for the individual to simply recall and repeat that same word.

Overall, a large amount of direct behavioral studies support differential processing in memory according to the modality and the goal of task (i.e., short-term or WM task). While neuro-imaging data do suggest that some areas of the brain are more active than others during these tasks, this evidence cannot be concluded as causal. Therefore, this review will function under the multi-modal assumption of WM. These systems appear to work in congruence to take in and update information.

Development of Working Memory

As demonstrated above, WM involves the communication between multiple brain areas. These areas have been argued to differentially develop across childhood with data supporting that "easier", lower-order abilities mature earlier than more higher-order processing (Best & Miller, 2010; Garon, Bryson, & Smith, 2008; Marsh, Gerber, & Peterson, 2008). That is, STM develops into domain specific WM and further into executive WM.

STM appears to develop gradually through infancy and early childhood. Numerous cross-sectional studies have demonstrated that an infant can successfully hold information by five to six months of age (Pelphrey & Reznick, 2002; Johnson, 2005; Reznick, Morrow, Goldman, & Snyder, 2004; Pelphrey et al., 2004). This is demonstrated through the use of delayed response tasks, where an infant is required to determine the location of a hidden toy after a delay. By 12 months of age, capacity increases are demonstrated in children in both longitudinal and cross-sectional studies; infants demonstrated the ability to hold the information (the hidden toy) for a longer period of time (approximately ten seconds) from a larger array of location choices (Pelphrey et al., 2004; Diamond & Doar, 1989). Studies utilizing verbal and object-spatial forward span tasks, have also demonstrated increases in memory capacity from age two through age five (Bull, Espy, & Senn, 2004; Ewing- Cobbs, Prasad, Landry, & Kramer, 2004; Gathercole et al., 1999).

More complex and higher-order aspects of WM, including updating and manipulating information, begin to develop during the preschool-age years. In visual WM, infants can successfully update the location of a prize under a set of cups around 15 months of age (Diamond et al., 1997). Alloway and colleagues (2004) demonstrated larger visual and verbal WM are present in children in four- to six-year-old children. A longitudinal study done by Gathercole and colleagues (2004) demonstrated that higher-order, executive component of WM can be adequately measured at age six years. Longitudinal and cross-sectional studies demonstrated a gradual increase in executive WM performance in children between 6 and 14 years of age (Brocki & Bohlin, 2004; Gathercole et al., 2004; Wu et al., 2011).

Some research has suggested a significant difference between the development of visual and verbal WM in children between the ages of 6 and 17 years (Conklin et al., 2007; Lambek & Shevlin, 2011). More specifically, when given a mix of verbal and spatial WM tasks, children tended to perform better with spatial information; however, both verbal and spatial development trajectories were similar. Lambek and Shevlin (2011) reported that the larger developmental factor with WM is the degree or level of processing; that is, mastery of a task is more dependent on the complexity of the task than the content of such task (Conklin et al., 2007; Lambek & Shevlin, 2011). For example, Gilchrist, Cowan, and Noveh-Benjamin (2009) demonstrated that "chunking" ability develops through childhood. In this study, seven-year-old children had greater difficulty holding multiple chunks of information than twelve-year-old children; however, there was no difference in the amount of words recalled in each chunk between groups. This suggests that the short-term capacity (number of words) does not change; however, the number of chunks or semantic units remembered increases, indicating the development of higher-order processing.

Neuro-imaging data support that the degree of processing informs development more than the mode of content presentation (i.e., verbal or visual), with more advanced, executive structures developing later in childhood. Evidence from functional and structural imaging studies indicate linear increases in fronto-parietal network activity, particularly in the pre-frontal cortex, in children between the ages of nine and 18 years. (Kwon, Reiss, and Menon, 2002; Nagy, Westberg, & Klingberg, 2004; Scherf, Sweeney, and Luna, 2006). Scherf, Sweeney, and Luna (2006) also demonstrated that increased lateralization and localization of activity in the brain during WM tasks. This indicates increases in the quality of activity, or more efficient processing, through adolescence and into adulthood.

Overall, there appears to be a gradual development of WM over time. STM develops within the first few years of life and, then, the more executive, WM abilities, involving manipulating and updating information begins to emerge during early school age years (i.e., five and six years of age). This continues to develop well into adolescence.

In order to differentiate WM from other important cognitive functions, a few important areas will be addressed in the following section. First, the impact of WM on language development will be discussed. There is also strong research base demonstrating WM's relationship to attention (Alloway, 2011; Martinussen et al., 2005; Raiker et al., 2011). Furthermore, WM has been categorized under a broader term for higher-order processing entitled Executive Functions (Barkley, 1997, 2001; Brocki & Bohlin, 2004; Latzman & Markon, 2009; Miyake et al., 2000; Wu et al., 2011).

Working Memory and Language

Speech and language development is related to WM. Some theorists have suggested that WM increases are solely a function of linguistic development (Ottem et al., 2007). They argue that language, alone, allows an individual to increase effective rehearsal and encoding ability. However, a study by Gilchrist, Cowan, and Naveh-Benjamin (2009) demonstrated that in a group of seven and twelve year old children, the ability to remember a set of sentences was based on the number of meaningful chunks of information and not amount of words in the sentences, with older children having a better ability to recall sentences with more chunks of information than younger children. In other words, memory span development accounted for the difference and not language development, suggesting that the memory system and language system are related but not interchangeable.

More likely, WM influences language development. STM has been linked to early language acquisition. Baddeley (1996) demonstrated the relationship between a child's ability to encode phonological information and the development of new words. Furthermore, Adams and Gathercole (1995) demonstrated that the verbal STM is related to the length of utterances in preschool-aged children. STM also significantly predicted a seven- and eight-year old children's ability to comprehend complex sentences (Botting & Conti-Ramsden, 2001).

An individual's ability to comprehend longer and more complex patterns of speech are also related to areas of WM. A study by Magimairaj and Montgomery (2012) demonstrated that WM ability, measured by a sentence recall task, significantly correlated with their performance on a standardized sentence comprehension task in a sample of 65 school-aged children (six years to twelve years); that is, lower WM ability predicted poorer performance on sentence comprehension. Verbal WM also significantly contributed to a child's ability to answer inferential questions on long oral narratives (Montgomery, Polunenko, & Marinellie, 2009).

Children identified with language impairments, such as individuals with aphasia or difficulty in the comprehension and production of language, are argued to have underlying WM difficulties. These children have more limited memory capacities than their typically developing peers, particularly verbal STM and WM (Archibald & Gathercole, 2006; Leonard et al., 2007). This affects their ability to retain language based information and impedes upon subsequent development.

In a study by Sung and colleagues (2009), a group of 20 participants with aphasia were administered verbal WM tasks, which were compared to their severity of symptoms.

Lower WM scores predicted higher severity in language difficulties. Past studies have also demonstrated that taxing or difficulties within the WM system will affect language processing (Miyake, Carpenter, & Just, 1994; Wright et al., 2007). Overall, WM is essential to many language-based tasks and contributes to language development.

Working Memory and Attention

Attention is the "gateway to conscious thought" (p.109); it is the ability to select a certain amount of sensory information for conscious processing (Hunt & Ellis, 2004). It filters out extraneous sensory stimuli from taking up cognitive energy, in order to efficiently interact with the environment. Attention is an important aspect inherent in all cognitive tasks (Baron, 2004). Across theoretical viewpoints of WM, there is agreement in the importance of attention; information must be attended to before it can be processed further. In Baddeley's (2000) model, the central executive (CE) area of WM is considered the attention system. It determines what information to focus on and where to direct it for further processing.

Early theorists believed that attention stemmed from an all or nothing selection process. Filter theories, named early- and late- selection theories, suggested that the attention system selects information before it reaches consciousness based on perceptual information (Broadbent (1958; Treisman, 1960; Deutsch & Deutsch, 1963; Norman, 1968). That is, more salient or meaningful information will be attended to while other information is not brought to conscious awareness. Over time, theories began to understand attention more as a limited capacity system (Hunts & Ellis, 2004). Later models, such as the capacity model of attention by Kahneman (1973) and the central bottleneck theory by Pashler (1984), argue that attention functions as a way to allocate resources based on a set of decision points to process information.

Most recently, neuropsychological data suggest that attentional processes can be divided into functional systems. Baron (2004) and Miller (2007) provide extensive support for differentiated sub-systems of attention; namely, selective/focused, sustained, divided, and alternating/shifting attention systems. Selective/focused attention describes the selectivity addressed by earlier attention models; it is an individual's ability scan an array of stimuli and consciously attend to a choice. Sustained attention describes the ability to maintain focus for a prolonged period of time. This is also referred to as vigilance. Divided attention pertains to the ability to respond to one than more stimulus or task at once. Shifting attention describes one's ability to juggle attention resources from one task to another.

An individual who has impaired attention also appears to have impaired WM. Researchers have demonstrated that in children with clinical attention difficulties, namely through a diagnosis of Attention Deficit Hyperactivity Disorder (ADHD), have great deficits in their WM. A meta-analysis of 26 articles investigating the relationship between WM and ADHD in children demonstrated moderate to strong effect sizes for spatial and verbal WM impairment in children with ADHD (Martinussen et al., 2005). Martinussen and colleagues (2005) found a stronger effect size in spatial WM over verbal WM; however, more recent research suggests that the level of deficit appears to be more related to the task demands than modality (Brocki et al., 2010). More specifically, in a longitudinal study by Brocki and colleagues (2010), in a group of 72 clinically referred five-year-old children, the degree to which WM was impaired predicted the severity of ADHD symptoms two years later.

Interestingly, these deficits do not appear to reflect all areas of memory in children with attention difficulties. In a study by Alloway (2011), a sample of 50 children diagnosed with ADHD had significant WM impairments; however, there was no significant relationship with STM tasks. Furthermore, Raiker and colleagues (2012), found that measures of WM mediated the level of impulsivity in children with a diagnosis of ADHD. This indicates that there is an intimate connection between executive attention and WM difficulties, as indicated in Baddeley's (2000) model. This relationship has also been demonstrated in non-clinical populations (Gathercole et al., 2008).

Working Memory and Executive Functions

Executive functions (EF) can be broadly defined as the regulation and maintenance of behavior to achieve optimal performance towards a future goal. It is considered the manager of cognitive processes necessary to make decisions and adapt in complex situations; in some literature, it is described as the "maestro" or "boss" of the brain (Hale & Fiorello, 2004).

Factor analysis studies have identified latent factors, or subdomains, that encompass EF ability. Those continually cited in the EF factor models are inhibition, WM, and flexibility (Barkley, 1997, 2001; Brocki & Bohlin, 2004; Latzman & Markon, 2009; Miyake et al., 2000; Wu et al., 2011). While some articles utilize the term "working memory" explicitly as a factor (Barkley, 1997, 2001; Brocki & Bohlin, 2004; Wu et al., 2011), others label the domain "monitoring" (Latzman & Markon, 2009) or "updating" (Miyake et al., 2000) However, all domain labels are operationally defined as the ability to work with and manipulate incoming information, and therefore represent WM.

Research has supported the similarity between WM and updating ability, as well as their differentiation from other executive function (EF) domains. A study by St. Clair-Thompson and Gathercole (2006) investigated the relationship between executive tasks and measures of WM. In this study, a sample of 51 English 11- and 12-year-old children completed two tasks measure shifting, where they needed to shift between counting items under a specific rule-set, two measures of updating, where participants needed to remember certain words or letters after the presentation of a group of variables, and two inhibition tasks, where participants were required to quickly inhibit a learned response in order to respond in a novel way (e.g., when seeing the word "red" in the color green and stating the color of the word). These EF measures were then compared to their performance on measures of WM, including a reverse verbal span task, a sentence recall task, and a visualspatial span task. Standardized measures of academic achievement in the areas of English, math and science were also administered. A significant relationship was found between all of the WM measures and the updating measures, as well as with the inhibition measures. Furthermore, a factor analysis revealed that the WM tasks and the updating tasks loaded onto the same factor, while inhibition loaded on a different factor and shifting did not significantly load on any measure. The WM and updating scores were also associated with a unique amount of variance in English and math scores; inhibition predicted a small amount of unique variance in tests of English, math, and science. These results are not surprising when the task demands of the measures differentially labeled as updating and WM are considered; both tasks require a respondent to hold and manipulate pieces of information internally to provide a new product. Therefore, one would expect a similar performance on updating and WM tasks, as well as how they relate to academic performance; however, WM ability still appears to be separate from other areas of executive functions, particularly inhibition.

Another study by St. Clair-Thompson (2011) compared 38 children with poor WM to 38 matched controls on measures of WM and executive functions. WM was measured by way of a counting recall task, where a participant is required to count the number of items in an array and then recall the successive tallies of each array. The same measures for inhibition and shifting described in the St. Clair-Thompson and Gathercole (2006) study were used; the updating measures were removed due to the strong evidence that WM and updating measured the same construct. In addition, a measure of planning, where individuals needed to recreate a block pattern under a certain rule-set, and a measure of attention, where participants were required to cross off pictures of a target object within a set of other stimuli on a sheet of paper, were completed. St. Clair-Thompson found that children with poor WM demonstrated significantly more difficulty on planning and attention tasks when compared to a group of matched controls. There was no relationship between WM and shifting or inhibition. Similar to previous discussions, this supports the relationship between attention, where lower WM is related to lower attention capacity. WM ability also continues to remain separate from other executive functions.

Working Memory and Academic Achievement

WM is inherent to our ability to gain knowledge and acquire new skills (Pickering, 2006). A strong link has been demonstrated between WM and academic achievement in school-aged children (Gathercole et al., 2004); that is, children with lower WM scores appear to also demonstrate lower performance on academic tasks. In a study by Alloway and colleagues (2009), a sample of 308 British children, ages five to eleven years, screened and identified with WM impairments, were administered tests of cognition, academic achievement, and behavioral rating scales. WM was assessed with measures of verbal STM, including forward digit and word span tasks, visual STM, including visual span tasks, verbal WM, including reverse digit span task and sentence recall task, and visual WM, including spatial span tasks. A standardized screener of general cognitive ability was administered; standardized measures of reading, including basic reading, reading comprehension, and spelling, and mathematics, including mathematical reasoning and numerical operations, were also given. Lastly, standardized behavior rating scales assessing attention difficulties and executive function difficulties were completed by teachers. A series of hierarchical regression analyses indicated that WM composite scores and the global general cognitive score accounted for a substantial amount of variance in the sample and both uniquely predicted academic performance. Teacher rating scales also reported higher behavioral problems in areas of attention and memory (i.e., “forgetting”). This supports WM ability as an independent indicator of academic attainment and behavioral difficulties in school-aged children.

Furthermore, Alloway and colleagues (2009) reported that older children appeared to perform significantly worse than the younger children on academic measures. This suggests that WM may increase the performance gap without intervention over time; however, due to the cross-sectional nature of this study, this conclusion cannot be reached with any certainty. These affects could be related to the increased WM demands in the classrooms of older children; that is, lower grade levels naturally provide more structure and support in lessons while higher grade level students are required to process larger and more complex pieces of information. Overall, WM difficulties independently predict children's academic attainment, with poorer academic performance seen in older children.

Gathercole and colleagues (2004) found a similar profile in a smaller sample of seven- (n = 40) and fourteen-year-old British children (n = 43). Participants were administered measures of verbal WM, including reverse digit span, sentence recall task, as well as measures of STM, including forward digit span, forward nonsense word task, and a word list matching task. These scores were then compared to their performance on the national curriculum assessments administered in the United Kingdom in English, mathematics, and science; these are similar to the standardized high-stakes testing seen in the United States. Both younger and older children's WM performance significantly correlated with all curriculum-based measures; that is, WM difficulties predicted lower scores on the national assessments in English, mathematics, and science. This, again, suggests a strong link between a child's WM ability and their performance on academic tasks. However, again, the cross-sectional nature of this study does not allow a definitive conclusion on the interplay between lower WM, age, and academic achievement.

A recent study by Taub and McGrew (2014) demonstrated this link in American children. In a large sample of two age-based stratified, age-based groups (i.e., 9- to 13-year-olds and 14- to 19-year-olds) who were administered a standardized cognitive and achievement battery, WM was found to have consistently strong links in all major areas of academic development.

WM difficulties also account for academic difficulties for children with clinical difficulties in attention and learning. Rogers and colleagues (2011) administered verbal and visual-spatial WM and STM tasks, as well as a standardized academic achievement battery to 145 clinically-referred Canadian adolescents ages 13 to 18 years with a ADHD diagnosis. The verbal WM tasks, namely a forward and reversed digit span and a letter-number sequencing task from a standardized intelligence battery, significantly predicted their composite reading score, including word reading, reading fluency, and reading comprehension measures, and their composite math scores, including calculation and word problems. Visual WM tasks, including spatial span, only significantly predicted participants' math composite score.

These relationships remained when attention difficulties were controlled for in the analysis, suggesting that WM and attention difficulties both play into academic performance in clinical populations. This is not surprising, considering the intimate link between attention and WM. However, the two verbal WM scores utilized in this study have been criticized for lack of purity in measurement of WM (Flanagan & Kaufman, 2004); that is, each task loads somewhat on simpler, STM which may have skewed results in this study.

WM difficulties are an integral aspect of learning difficulties in children with learning disabilities. Swanson (1993) identified a group of 123 clinically-referred children with either identified with a learning disability in reading or math, identified as an achievement-matched participant (i.e., the participant had low achievement but no learning disability), or identified with average achievement ability were administered verbal and visual STM and WM tasks. Results indicated a non-significant difference between participants with reading and math learning disabilities, suggesting that there is a global deficit in WM when identified with a learning disability. There was a significant difference between the learning disabled group's WM scores and the matched average ability group. Interestingly, when the learning disabled groups were compared to the achievement-matched group, stronger differences were seen in the STM tasks than in the WM tasks. This suggests that as processing demands increase in the memory system, performance gaps between low performers and those with a learning disability decreases. Therefore, low performing students, no matter the phenotype, seem to struggle with WM.

An individual’s WM ability is strongly linked to development of basic and higher-level reading skills. Nevo and Breznitz (2013) investigated the relationship between WM and reading development in elementary school students. A sample of 97 Israeli kindergarten students were administered tests of reading and WM. For reading, three informal individually administered decoding and reading fluency measures and three group administered reading comprehension tasks were given. Verbal and visual STM and WM tasks were also administered from a standardized WM battery; however, all verbal WM measures utilized were converted to Hebrew, which poses a problem for standardization. They found strong correlations between verbal and spatial WM and all reading skills measured. Furthermore, when grouped according to decoding scores, poorer decoders received significantly lower scores on verbal WM measures. After a one-year follow-up, only the relationship with verbal WM tasks persisted. This indicates that earlier on, that WM globally impacts reading skills; however, over time, verbal WM difficulties persist. However, these results cannot necessarily be generalized to English-speaking children.

Compton and colleagues (2012) demonstrated this link in American children. They followed 684 randomly-selected students from third to fifth grade and assessed them annually on standardized academic measures, including word reading, reading comprehension, calculation and math problem solving. These were compared to their scores on standardized measures of cognition, including non-verbal problem solving, processing speed, concept formation , standardized measures of language including grammatic closure, listening comprehension, and vocabulary, and standardized WM measures, including sentence recall task, and a reversed digit span task, collected at the beginning of third grade. Children identified with a word reading learning disability, by way of inadequate progress in academic areas plus cognitive weaknesses, had persistently more difficulties in WM than children without a word reading learning disability. However, no other learning disability subgroup was found to have significantly lower WM difficulties.

This somewhat unexpected result may be due to testing procedures in the study. While participants were repeatedly administered academic measures in the study, the WM and cognitive variables were only collected at Time 1. The scores during a single administration may not have been the most representative of their abilities. Furthermore, the authors admit that these results may have been due to low samples sizes for the learning disabled groups compared to the non-learning disabled group when performing secondary multivariate analyses.

Swanson and O'Connor (2009) investigated the influence of reading fluency interventions on WM and reading comprehension. In this study, a sample of 155 second and fourth grade American students were screened for poor reading skills using a receptive language measure and were randomly assigned into one of three reading intervention conditions for 20 weeks: repeated reading, continuous reading, or a control group. Students in the treatment condition either received 15 minutes of reading a page of text three times (i.e., repeated condition) or 15 minutes where students read multiple pages without stopping (i.e., continuous condition) three times per week. The students in the control condition received no additional intervention. Students were measured at three times on standardized measures of receptive vocabulary, word reading task, reading rate and comprehension task, and a sentence recall task to assess verbal WM. Overall, students in the continuous reading condition demonstrated higher fluency increases than in the repeated reading or control condition. Specific to WM results, they found that verbal WM predicted post-test reading comprehension ability in all conditions, independent of reading fluency improvement. That is, no matter the intervention, WM alone predicted reading achievement. In addition, hierarchical linear modeling that partialed out pre-test scores, vocabulary, and word-attack skills demonstrated that WM significantly related to posttest outcomes in text comprehension. However, no significant interactions arose between intervention groups and WM ability on post-test measures of comprehension, vocabulary and fluency. This demonstrates that WM has a specific influence on reading development.

Interestingly, there is mixed support for a relationship between WM and reading comprehension. Some studies have demonstrated a strong link between WM and sentence or text-level comprehension (Berninger et al., 2010) and words (De Beni et al., 1998). Berninger and colleagues' (2010) longitudinal study evaluated WM's contribution to reading writing outcomes in urban American students; second (cohort 1 $N = 122$), fourth (cohort 1 N = 114; cohort 2 N = 108), and sixth (cohort 2 N = 105) grade students were utilized for this article. Participants were administered three different span tasks, two sentence recall tasks, standardized reading tasks (i.e., word reading, reading comprehension), and standardized writing tasks (i.e., handwriting, spelling, & written expression). They found the spans tasks, collectively, contributed unique variance in reading and writing skills for students in all grade levels; it explained all reading and writing outcomes for second graders, it predicted word reading, handwriting, spelling, and written expression in fourth graders, and predicted word reading and spelling in sixth graders. Sentence recall tasks contributed unique variance to a lesser extent; it did not uniquely predict any measures in the second grade cohort, it predicted reading comprehension in the fourth and sixth grade cohorts. WM measures did not predict handwriting or written expression at the sixth grade level. This clearly demarcates WM as pertinent to the developing student; it continually predicts reading and writing outcomes across grades. Interestingly, it also outlines how WM develops and differentially contributes to relevant academic abilities over time; in second grade, where most students are still reading through lower level tasks, such as spelling and word reading, which were significantly predicted by WM whereas reading comprehension was predicted by WM in older, sixth grade students.

De Beni and colleagues (1998) demonstrated the relationship between reading comprehension and WM in non-English speaking undergraduate students. A group of 44 undergraduate students studying at university in Italy were initially administered a reading comprehension task where they read five texts and asked to answer four multiple questions on each text; it is not clearly stated if the participants were able to access the passages while answering questions. Based on their scores, the researchers divided the students into two groups, good and poor comprehenders; these groupings were confirmed to be significantly different from each other. Then, students were individually administered an Italian version of a sentence recall task and a forward and reversed digit span task. It was demonstrated that the poorer comprehending group performed significantly worse on a working memory task. More specifically, while this group performed similar to the good comprehenders in recalling a span of words, when asked to reposition the words in a novel order, they demonstrated significantly more difficulty. This was replicated in a second experiment with a different type of listening span task and demonstrates that comprehension difficulties is found in non-English speaking populations.

Garcia-Madruga and colleagues (2013) also found a relationship between reading comprehension and WM measures in a sample of 31 eight and nine year old Spanish children that were provided a reading comprehension intervention. In this study, students were randomly assigned to the reading intervention group or a control group and administered pre/post testing in reading comprehension, WM and nonverbal fluid reasoning. While the intervention itself did not demonstrate effectiveness between groups, the pre- and post- testing data demonstrated strong positive correlations between an individual's ability to answer multiple choice questions on a multi-paragraph reading passage *without* referencing the passage and a listening span WM task; interestingly, there was no significant correlation with nonverbal fluid reasoning.

WM was unable to predict performance of the ability of student’s to answer multiple choice questions on longer reading passages in a sample of 585 third, seventh, and tenth grade American students (Kershaw & Schatschneider, 2012). In this study, students' oral reading fluency, listening and reading comprehension, working memory, verbal and non-verbal reasoning and decoding abilities were assessed on three different occasions. Utilizing a latent variable modeling approach, the authors found that beyond decoding and language comprehension, fluency was the only variable that significantly predicted reading across all three grade levels; working memory did not predict performance. This is not surprising, as in this study, the participants were able to access the text while responding respond to multiple choice questions, which should not significantly tax WM resources and demonstrate differential performance.

The discrepancies between these studies is related to the methods used to measure reading comprehension; the task demands in reading comprehension by Kershaw and Schatschneider (2012) did not adequately tax WM. In the reading comprehension tasks utilized, one must read a passage and use the text to answer questions. There is less of a need to hold and maintain information in one's mind if the passage is present to refer back to, making WM less necessary to complete the task. Most reading comprehension tasks also utilize multiple choice responding, allowing the respondent to use process of elimination or context clues for answers. In addition, text comprehension questions may have asked more literal or factual questions and not higher-order inferential questions, which would also put less demand on WM.

This relationship has been demonstrated in a study by Chrysochoou and colleagues (2013). A sample of 92 Greek eight and nine year old students were administered short-term verbal memory tasks, namely measures where the participant was asked to recall digits and words, and WM tasks, namely listening span and backward digit recall, as well as a measure of reading fluency and comprehension. Results indicated that working memory measures, with the listening span task having the strongest relationship, were significantly correlated with the reading comprehension task while the short-term memory measures were not related to comprehension. In further analyses, they also demonstrated that WM was highly related to the participants' ability to inferential comprehension questions; that is, participants' with higher working memory were better able to answer questions that required connections to be made with the text and outside knowledge. Therefore, it appears that WM is necessary for the higher-order thinking involved with more inferential and generative questions in reading passages; tasks that do not tap into this executive area will not show a relationship with WM tasks.

In the area of mathematics, researchers have found that WM moderates problem solving ability. Zheng, Swanson, and Marcoulides (2011) administered measures of WM, mathematical computation, word problem-solving, and word recognition to a sample of 310 second, third, and fourth grade American elementary school children. They found that all components of WM (i.e., phonological loop, visual-spatial sketchpad, and executive system) predicted problem-solving accuracy; that is, lower WM predicted lower problem-solving ability. Furthermore, reading ability and computational proficiency mediated the effects of WM. This study demonstrates that global WM affects problem-solving ability; however, academic knowledge (e.g., reading ability and basic mathematical ability) can also influence this relationship. This would suggest that the relationship between WM and certain academic abilities are not as straightforward as previously postulated.

Swanson (2011) also found longitudinal effects of WM on problem-solving accuracy. In this study, a sample of 127 American first grade children were administered tests of cognition, WM, mathematics, reading, naming speed, and inhibition over a span of three years. Swanson (2011) found verbal WM tasks predicted problem-solving accuracy in the third grade and concluded that these components are crucial in word problem-solving development.

WM is related to writing and note-taking ability. Olive and Passerault (2012) argue that visual spatial WM is inherent in composing text. Piolat, Olive and Kellogg (2005) describes note taking as a dynamic process that utilizes a great amount of cognitive resources to produce a written record of information. In particular, the authors highlight the purpose of note-taking and the speed of processing, or WM, necessary for comprehension and production. They propose a dual and triple-task method for measuring cognitive effort during note-taking and the importance of understanding the differential cognitive loading when engaging in diverse tasks (i.e., notetaking from lecture vs. notetaking while reading). Another article by Peverly (2006) also argues that writing fluency implies automaticity of skill and frees up WM capacity for critical thinking. He concludes that within the context of writing, instruction should focus on fluent writing skills to allow for more efficient and productive use of planning, monitoring, and revising skills.

In a two experiment study by Olive, Kellog, and Piolat (2008), 72 undergraduate students were required to complete WM tasks, namely by way of a listening or spatial span task, while constructing an argumentative essay. Writing fluency significantly decreased with the presentation of visual and verbal WM tasks, indicating that WM is taxed during the writing process.

A study by Kellog, Olive, and Piolat (2007) investigated the connection between working memory and written language production. Across two experiments, a sample of 60 and 80 undergraduate students were asked to define vocabulary words (i.e., concrete or abstract nouns) while also completing a working memory task (i.e., verbal working memory or visual working memory). They found that the groups asked to complete a verbal working memory tasks had significantly more difficulty writing definitions of both concrete and abstract words.

Overall, studies provide support for a direct relationship between WM and specific academic difficulties. Poor WM has been shown to negatively influence basic reading, reading comprehension, and mathematical problem-solving. It is also suggested that these difficulties worsen over time.

Bottleneck/Overload Hypothesis

One hypothesis, coined the "bottleneck" hypothesis, postulates that academic difficulties in children with lower WM are a result of lost information during classroom-based instruction. Alloway and Gathercole (2006) argue that WM influences a child's ability to attend to and perform all instructional-based tasks, which impacts their ability to retrieve and use information learned. In children with lower WM, children become overloaded with information and miss out on new knowledge after a capacity limit is reached. For example, students with lower WM may fail to follow multi-step directions in tasks, such as mathematical equations. When attempting a problem, students with lower WM cannot hold all the steps necessary to solve that equation and, therefore, make careless errors, cannot complete the example, and ultimately, miss out on new pieces of knowledge. If this happens continuously, where more basic skills are missed, later or more advanced knowledge will not be supported and performance diminishes.

In order to assess support for this theory, Gathercole and colleagues (2008) created a two-part study investigating the effects of WM on five- and six-year-old children's ability to recall and execute classroom-like tasks. In Study 1, a sample of 100 British children with varying WM abilities were tested on orally presented sentence and poem recall tasks. As predicted, children with lower WM had significantly lower scores on both recall tasks. In Study 2, these same participants were required to recall and carry out increasingly difficult oral instructions. The results of this second study indicated that the children in the lower WM group were less accurate in carrying out instructions; however, they did not significantly differ from the average WM group in immediate recall. Overall, it appears that children with below average WM abilities have more difficulty manipulating information and carrying out instructions during classroom-like activities, supporting the "bottleneck" hypothesis.

Working Memory Interventions

With the demonstration of instructional-based difficulties in these children, studies have begun to explore how altering the instructional environment may influence performance in children with lower WM. They hypothesized the use of instructional modification techniques, such as chunking or repetition of instructions, may remediate academic difficulties in children with lower WM, under the assumption that lowering the cognitive demands of instruction on these children will increase academic achievement. In a study by Elliot and colleagues (2010), a sample of 256 British children, ages five to eleven years with lower WM, were assigned to one of three conditions, a teacher-facilitated WM intervention, a direct-teaching intervention, or a control group. In the WM intervention group, teachers were instructed to utilize hypothesized supports for WM, such as chunking instructions or checking for understanding; in the direct-teaching intervention, teachers were instructed to utilize focused, subject-related skill development, such as reviewing addition or subtraction rules. Results indicated no significant differences on cognitive or academic measures in either intervention group compared to the control. The authors postulated that the indirect nature of this intervention did not lead to concrete improvement in WM or academic performance.

However, extraneous factors may have influenced the results in this study. First, it was reported that each school only participated in one condition, creating a higher likelihood for environment to affect the results. School climate may have disproportionately influenced the conditions, causing a misrepresentation of effectiveness; for example, a more positive or negative school climate, and not the condition, can cause an increase or decrease in a student's performance. Furthermore, the authors appear to be using the same sample of children for a number of studies, some of which have already reviewed (i.e. Alloway et al., 2009). Testing effects on participants may have influenced results; that is, children who continually are tested on similar measures, like digit span-like tasks, may improve due to practice and not an intervention. Therefore, new sampling and uniform testing conditions should be attempted in future research.

Other studies have looked into interventions that directly remediated WM in children with hopes of generalizing improvements to academic performance. Some researchers attempted remediation through WM strategy training, where children were taught techniques, such as rehearsal or visual imagery. In one such study, St. Clair-Thompson and colleagues (2010) investigated how a working memory strategy training program may influence a child's performance on memory and academic tasks. In this study, a sample of 254 British children between the ages of five and eight (*M* age = 6 yrs 11 mos.) were assigned into one of two conditions based on existing classroom structure: a *Memory Booster* training classroom or a control classroom. All participants were administered working memory tasks (i.e. reversed digit span, spatial span, and sentence recall) before and after a 6-8 week intervention; "subgroups" of the sample were also administered tasks in following instructions (N = 77), mental arithmetic (N = 34), reading (N = 81) and mathematics (N = 81). Results indicated a significantly greater improvement in scores on the digit and sentence recall tasks for children in the *Memory Booster* condition; the intervention group also demonstrated greater improvement in the following instructions and mental arithmetic tasks. The authors concluded that this computerized training program can improve a child's working memory capacity in the classroom; however, they state that the program's success appears to remain specific to the area of working memory and does not generalize to academic improvement. Besides the fact that British population of children may not necessarily generalize to American children, another question remaining is how the authors specified criteria for the "subgroups". If the "subgroup" selection was based on a variable like volunteerism, motivation may be a cause for significant improvement; more conscientious students may be more diligent in practicing and applying aspects of the program, allowing better generalization of skills and higher scores. On a different note, if the "subgroup" was disproportionate in ethnicity, which is unclear since all demographic information is absent from the article, the improvement may be a result of race, socio-economic status, or something else entirely. This leaves the positive results of the *Memory Booster* program on classroom behavior scores up for interpretation. There is a need for explicit explanations of demographic information in order to generalize results to the appropriate population and for replication in future studies.

Other researchers have attempted direct WM remediation through repeated exposure to recall tasks, also known as drilling. Loosli and colleagues (2012) investigated the effectiveness of a computerized working memory training program on cognitive and academic outcomes of third grade children. Forty Swiss children (*M* age = 9 yrs 9 mos.) participated in this study; 20 were administered a 2- week computerized working memory training and 20 served as matched controls by gender, age and pretest *Gf* scores. Participants were pre- and post-tested on *Gf* (fluid reasoning) and reading ability (only decoding and fluency). Results indicated that the trained group demonstrated significant increased in "working memory" and that the trained group performed significantly better on the reading fluency task than did the matched controls after the intervention. The authors determined that working memory training can improve reading scores in school-aged children; they also stated that these results may indicate improvement in working memory and attentional control in the experimental group.

However, as a quasi-experimental design, clear conclusions cannot be drawn; matched control groups cannot account for many extraneous variables. First, the controls were pulled from different classrooms while the experimental group participants were all from the same classroom; it cannot be ascertained that the improved reading scores were due to the training itself or to the possibility of environmental/instructional improving scores, such as a warmer, more attentive teacher. It is also peculiar that this study did not test for working memory ability pre- and post-intervention. It would be important to use a standardized measure of working memory ability to for baseline data to gauge improvement and assess intra- and inter-group differences. Without it, there is no way to conclude that there was true change in working memory. This leads to the last point of internal validity; when assessing the effectiveness of a new training program, it would be important to see if the tasks' content/construct validity in working memory and attention to tease out what improvement in training truly is indicating.

In another study, Holmes, Gathercole, and Dunning (2009) administered 42 eight to eleven year old British children with lower WM an adaptive or non-adaptive version of a five to seven week long WM training program; improvement was gauged by performance on cognitive and WM tasks. Results indicated that the adaptive group demonstrated significantly greater gains in verbal, visual spatial, and applied WM tasks than the non-adaptive group. Furthermore, a significantly greater number of participants in the adaptive group increased their WM scores to fall within the average normative range after the intervention. Mathematical reasoning gains were also seen in the adaptive group at the six month follow-up. The authors concluded that an adaptive model of WM training is superior to more standard computerized programs. The authors also concluded that an adaptive model of working memory training is superior to more standard computerized programs, such as Cogmed (Klingberg et al., 2005). As stated in other article critiques, it will be difficult to generalize these data to American children due to the many environmental variables that may play a part in this study. Furthermore, it was reported that each school only participated in one condition, creating a higher likelihood for environment to affect the results. School climate may have disproportionately influenced the conditions, causing a misrepresentation of effectiveness; for example, a more positive or negative school climate, and not the condition, can cause an increase or decrease in a student's performance.

A study by Chein and Morrison (2010) demonstrated benefits of WM training benefits in an American population. A group of 42 American undergraduate students completed four weeks of an adaptive WM training program. At its completion, students increased their WM capacity. Further, students demonstrated increased performance on other tasks, including a measure of executive attention, and reading comprehension. This provides the most promising support that training WM allows a transfer of skill to other avenues. However, the population utilized was undergraduate students and may not generalize to school-aged children.

Other studies utilizing similar adaptive computerized programs have shown some success in clinical populations. Klingberg and colleagues (2005) found significant reductions in inhibition and inattention symptoms in a sample of 53 seven- to eleven-year-old children diagnosed with ADHD. A study by Holmes and colleagues (2009) also demonstrated improvement in visuo-spatial WM in 25 English children diagnosed with ADHD compared to a medication-only control group. A study by Dahlin (2010) demonstrated increases in reading comprehension and WM after five weeks of the computerized WM program in a sample of 57 Swedish children with learning difficulties, ages nine to twelve years of age.

Overall, these studies hold the most promise to the future of WM interventions, as they highlight a few unique variables that are lacking in other studies, that is, individualized remediation and follow-up testing. First, an adaptive model to intervention allows flexibility in structure; it can take into account the severity of WM deficits and allows children to work at their own pace. These variables may be the key to effective WM remediation.

In addition, Holmes and colleagues' (2009) study demonstrates the need for longitudinal data on these training program. While most other articles used a pre/post-test design and found limited significant academic improvement, this study assessed academic progress at a six month follow-up and found significant change. It may be that academic improvement in lower WM students progresses slowly and takes time for academic gains to be seen.

Assessment of Working Memory

WM is generally assessed through individually administered "span tasks" (Gathercole & Alloway, 2008; Pickering, 2006). Levin, Thurman, and Kipert (2010) indicate that 11 of 15 measures that are largely used for WM assessment are considered span tasks. They are reportedly the most utilized assessments in psychology and are widely considered reliable and valid (Conway et al., 2005). Leffard and colleagues (2006) indicate that WM span subtests are largely from standardized cognitive batteries, such as Digit Span and Letter-Number Sequencing subtests from the Wechsler Intelligence Scale for Children – Fourth Edition (WISC-IV; Weschler, 2003). These tasks require the individual to recall and transform an increasing number of sequenced items, such as letters and numbers.

Although popular, span tasks have recently come under scrutiny for lacking the ability to indicate WM behavior in real world settings (Levin, Thurman, and Kiepert, 2010). This is considered ecological validity; it is a measure's ability to predict functioning in the everyday environment. Brunswik (1956) and Bronfenbrenner (1979) began the argument stating that there was an overgeneralization of results from lab-based tests to real-world situations. Recently, cognitive psychologists and neuropsychologists have argued that training behavior in settings that mirror the applied settings is crucial to the success of any intervention put into place (Savage et al., 2007; Spooner & Pachana, 2006; Thurman & Kiepert, 2008).

Ecological validity is composed of two main areas, namely verisimilitude and veridicality (Franzen & Wilhelm, 1996). Verisimilitude references a measures' ability to match the task demands of the construct in the everyday environment. A measure with good verisimilitude should mirror tasks in real-world settings. Veridicality references the relationship between the measure and traditional assessments of the construct. This also can be considered convergent and divergent validity.

Existing assessment of WM lacks ecological validity. Span tasks lack verisimilitude; in other words, a task requiring a child to repeat a span of non-meaningful items, such as letters or numbers, is not a necessary skill within the classroom (Levin, Thurman, & Kiepert, 2010). This makes identifying and understanding WM difficulty in the classroom difficult, particularly for individuals who have less knowledge in cognitive skills such as teachers and parents. Span tasks also do not inform how to intervene on WM difficulties in the classroom; they provide limited information into how to support development of WM in children.

Levin, Thurman and Kiepert (2010) suggest two main methods of acquiring ecologically valid WM. First, WM could be assessed by way of behavioral data in their natural environment; that is, what WM manifests as, behaviorally, when approaching a demanding task. Behavioral rating scales, filled out by the individual or other informed adults such as parents and teachers, are considered adequate at measuring overt behavior. In these measures, teachers and/or parents rate the level of specific behaviors in a given child over a period of time and, then, compare the level of behavior to a normative or clinical population.

While behavioral rating scales can be very informative, researchers have criticized these forms as poor indicators of the level of functioning on more cognitive, internal states. Some research has indicated that it is difficult to measure more internal states behaviorally (Kolko & Kazdin, 1993). In this study, Kolko and Kazdin (1993) compared the ratings of parents, teachers, and the children in clinical and non-clinical samples on measures of internalizing and externalizing behavior. They found differences in agreement across raters according to behavior type; more specifically, there was less agreement across respondents in internalizing behaviors than externalizing behaviors. This sheds some doubt on the ability of behavioral scales in assessing internal cognitions and may reflect adults' difficulty in perceiving internal states in children.

Furthermore, behavioral measures with specific WM subscales have also come under question. Normand and Tannock (2012) argue that one of the most well-known measures of WM, the Behavior Rating Inventory of Executive Functions (BRIEF; Gioia, Isquith, Guy, & Kenworthy, 2000) does not adequately assess WM. They state that the 86-item rating scale measures various aspects of executive functions, including planning, inhibition, shifting, emotional control, monitoring, and organization, and overlap with many features of ADHD. Therefore, WM difficulty may be confounded by attention difficulty.

Other research demonstrates that the BRIEF lacks convergent validity. First, Gioia and colleagues (2000) do not indicate any data suggesting validation of the WM subscale to other performance measures of WM. Mahone and colleagues (2009) also demonstrated parental ratings of WM on the BRIEF do not correlate with either span task measure in a population of 35 typically developing children ages 5-17 years. Similarly, Forchelli, Carris, and Fiorello (unpublished manuscript) did not find a significant relationship between parent and teacher rating scales and WM measures in a sample of 90 clinically referred children ages 5-18 years of age.

More recently, researchers have attempted to develop rating scales devoted to WM measurement. The Working Memory Rating Scale (WMRS; Alloway, Gathercole, & Kirkwood, 2008) is the only published WM scale to date (Levick, 2011). Alloway, Gathercole, Kirkwood, and Elliot (2009) provided preliminary evidence suggesting good reliability for the 20-item measure. It also demonstrated significant relationships with performance measures of WM, including Digit Span from the WISC-IV in a sample of 417 English children, indicating good validity. However, replications by non-biased researchers, namely Normand and Tannock (2012), indicated that the 20-item WRMS had a poor confirmatory factor analysis fit in a sample of 524 Canadian, six-to nine-year-old children. Therefore, behavioral rating scales appear to lack the ability to adequately measure WM behavior.

The second type of ecologically valid measurement suggested by Levin, Thurman, and Kiepert (2010) was through a performance on a classroom-based task that utilizes WM, such as requiring a child to follow a set of instructions. Studies have recently begun to utilize these classroom-based tasks to measure WM ability in the classroom. As discussed previously, Gathercole and colleagues (2007) utilized a following classroom directions task requiring children to perform simple actions with school supplies, such as "touch the white bag then pick up the yellow ruler then put it in the blue folder" (p. 1027). This task was found to be significantly related to other span-like WM tasks. Similarly, St. Clair-Thompson and colleagues (2010) utilized a following directions task to assess change in a WM training program. This measure required students to follow increasingly difficult directions, such as "point to the picture at the top of Page 3 and copy it twice" (p. 208).

A study by Forchelli, Cariss, and Fiorello (unpublished manuscript) demonstrated support for the veridicality of a following directions task. In a sample of 92 clinically-referred urban American students, standard span tasks were strongly correlated with a following directions task. This task was also related to areas of academic achievement, including listening comprehension, reading comprehension, math computation, math problem solving, and writing composition.

Overall, it appears that ecologically valid measures are important to help inform remediation of WM difficulties. Two approaches discussed were behavioral rating scales and performance based measures. Behavior rating scales appear to be in their infancy in tapping aspects of WM; most do not adequately represent WM ability. Applied, performance-based tasks appear to have more promise, demonstrating support in verisimilitude and veridicality.

External Validity of Working Memory Measurement

In addition to the necessity for development of more ecologically valid measures, there is also a need to enhance the external validity of WM assessment and research. The majority of researchers developing measures and remediation programs in the literature are from Canada and the United Kingdom. They have also been criticized for utilizing the same sample to develop measures and implement interventions (Normand & Tannock, 2012).

It is important to be able to generalize findings and measurement to children within the American educational system. There has been a shift in the approach to identifying children with academic need from traditional "test and place" to Response to Intervention (RtI) (Brown-Chidsey and Steege, 2005). They describe a system where students are continually assessed and monitored in their academic growth compared to national benchmarks. If students do not meet these standards, they are provided increasingly individualized supports to address these needs. This model to assessment and intervention has demonstrated strong, long-term improvement in students' academic achievement (Case et al., 2010; Speece et al., 2003).

Notwithstanding, the RtI framework does not address identification or intervention for WM development/need. It would be important to adopt a method to screen and assess difficulty in this area due to overwhelming support to WM's impact on the developing learner. As mentioned previously, Taub and McGrew (2014) linked WM and achievement in American children and argue that WM is a major component to "cognitive efficiency," which mediates cognitive skill and development. Due to this overwhelming connection, they suggest the utilization WM ability as a screener to "precursors" of more serious learning difficulties. However, WM measures reviewed are all individually administered tasks, which are both time and resource intensive; no screener exists to put this empirically-supported claim to practice. Researchers have suggested the use of school- or classroom-based screeners as an efficient tool in identifying behavioral difficulties (Farmer, Burns, Phillips, Angold, & Costello, 2003). A short, quickly administered tasks is needed.

Group administration of WM skill would facilitate earlier identification. Many researchers have argued that early intervention facilitates skill development in educational and social realms is desirable (Heckman, 2006; Sonuga-Barke, Koerting, Smith, McCann, & Thompson, 2011). These researchers argue that cognition develops hierarchically and interdependently and influences later performance, making it ideal to intervene when the brain appears more plastic and malleable to change. Wass and colleagues (2012) support this argument, suggesting that early WM intervention is related to a widespread transfer of skills. In a meta-analysis of cognitive training programs, mainly ones geared to enhance WM or attention capacity, they found a small effect size for studies targeting younger students leading to significantly more improvements on standardized cognitive and academic measures; studies only considering typically developing children had a moderate effect size.

This transfers well to an argument for early intervention on WM difficulties. The “overload hypothesis”, postulated by Gathercole and Alloway (2008), references a difficulty in taking in continuous information, which compounds subsequent processing.

This is also supported by previously reviewed research stating that WM deficits make academic progress increasingly difficulty over time (Alloway et al., 2009). Therefore, identification of WM difficulties is imperative in these communities to address differential learning needs in schools.

Test Construction of Psychological Tests

Validity concerns the ability to make certain interpretations with performance on a specified measure by way of relationships with outside evidence and support (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2014). It is the careful attention given to possible distortions of meaning in data. More historical models of validity have divided it into three areas, mainly content, criterion, and construct validity (Cronbach & Meehl, 1955). However, Messick (1995) shifted this view by encapsulating validity into six aspects, including, content, substansive, structure, generalizability, external factors, and consequential. In the AERA, APA, and NCME (2014) standards of validity, they were further divided into five types of evidence: internal structure, test content, response processes, relations to other variables, and consequences of testing. In order to validate a test, information from multiple types of evidence should be considered in conjunction with other supportive data.

In order to provide evidence of internal structure, internal consistency or internal reliability is measured. Cronbach (1988) discusses the use of the alpha coefficient to assess the consistency of responses of a certain set of items in a measure, also termed Cronbach's Alpha. It determines a measure's ability to assess a unitary construct.

Test content is equivalent to content validity. This outlines a tests ability to encompass all areas of a given construct; it is "the specification of the boundaries of the construct domain to be assessed (p. 745, Messick, 1995)." Messick argues that this can be done by way of analysis of the tasks within the assessment. He also mentions Brunswick's (1956) use of "ecological sampling" mentioned previously in this review under ecological validity, where functional importance of the task is assessed. These are traditionally judged by professionals in the field and are supported by past empirical documentation.

Response processes refers to the fit between the definition of the construct and the nature of performance or response requested (AERA, APA, & NCME, 2014). It is considered the "evidence of data integrity such that all sources of error associated with test administration are controlled or eliminated to the maximum extent" (p.38, Downing, 2003). This involves investigation of both the cognitive output and processes of the test taker, as well as the adherence to scoring criteria by test scorers. For example, not only should consideration be given to what the individual needs to utilize to respond to content but how that response is, then, evaluated.

Relation to other variables represents criterion-related validity. In order to provide evidence of a test's relationship to other variables, test scores are correlated to existing measures with good validity and reliability (Cronbach & Meehl, 1955). This is done by way of convergent and divergent validity. Convergent validity represents how well a measure agrees with other existing measures deemed appropriate to measure the construct. Divergent validity is considered how well the measure can be distinguished from other measures not related to a construct. Campbell and Fiske (1959) argue that convergent validity correlations should be significant with similar tasks and statistically higher than correlations with dissimilar constructs. Further, AERA, APA, & NCME (2014) indicate a need to investigate how well the score from the measure predicts performance. This involves predictive and concurrent validity. That is, its ability to predict future or current performance in a related area.

Consequences of testing or consequential validity refers to the external impacts of test administration and scores to a participants' functioning (AERA, APA, & NCME, 2014). This type of validity intends to begin to address more long-term outcomes of a measure. Downing (2003) reports that variables that are necessary to consider for this area of validity include the effect a passing rate on a measure may have on subsequent decisions (such as receiving academic services), and the subjective effects of receiving such grade.

Overall, AERA, APA, & NCME (2014) articulate that a test will meet the standards of validity if there is "clear articulation of each intended test score interpretation for a specified use," (p. 23) with ample evidence provided to support each intended interpretation. This should include its utility and appropriateness in varied populations and settings. For the purpose of this study, only a few areas of validity are addressed. As this study focuses on the development of a measure, more emphasis is placed on the measure itself. This would suggest a need to investigate internal structure, test content, and response process validity. In addition, in order to investigate larger construct validity, the measures relation to other variables was investigated.

CHAPTER THREE

METHODOLOGY

Sample

Elementary school classrooms spanning kindergarten to third grade classrooms were recruited from 19 different school districts in the greater Philadelphia area by way of email and phone calls to school administration. A total of five schools agreed to participate in the study, with one to two classrooms per school district. Table 3.1 provides overall demographic information for the participating schools. Data on location in reference to a major city (i.e., urban, suburban, or rural), public or private funding, number of students enrolled, percentage of majority ethnic categories, and Title 1 funding (i.e., free or reduced lunch) were collected from school administrators and public online databases (ElementarySchools.org, 2014)

Table 3.1

Location of school, school funding source, number of students enrolled, student ethnicity, and percentage of students requiring free/reduced lunch

School Name	Location	School Funding	Enrollment	Ethnicity	Free/Reduced Lunch Ratio
School 1	Urban/ Suburban	Religious Affiliated	449	58% Caucasian	N/A
School 2	Urban/ Suburban	Charter	2147	1% Caucasian	$\geq$ 40% Free/ Reduced Lunch

Table 3.1, continued

School 3	Urban/ Suburban	Religious Affiliated	86	27% Caucasian	N/A
School 4	Urban/ Suburban	Charter	1284	81% Caucasian	N/A
School 5	Urban/ Suburban	Public	451	79% Caucasian	2% Free/Reduced Lunch

Classroom participation was decided upon by the principal of each school and teachers provided consent to participate in the study before student recruitment was initiated (see Appendix C). Parents of the participants provided written consent and participating children also provided verbal assent (see Appendix B). Teachers sent home consent packets to parents of all students within the classroom to complete. These packets included an introduction letter signed by the researcher and the principal of the participating school, a demographic form, and a BRIEF parent rating scale (see Appendix A). The introductory letter requested that completed packets be returned with the student and collected by the teacher. In an attempt to increase participation, one week after the initial distribution, a reminder by way of a colored paper slip was given to students who had not returned the packets. A total of two weeks passed between distribution and final count of participants. The return rate ranged from 20% of packets returned from School 3 to 86% of packets returned from School 4.

Both regular education students and students receiving special education services were included in the study. Inclusionary criteria was based on their current education placement. That is, children whose educational needs are met in a regular education or inclusive classrooms were included; students who were placed in a more specialized setting were planned to not be included in the study. Students who were being evaluated for special education services during the study were also planned to not be excluded from the study in order to not influence psychoeducational testing performance. Notwithstanding, no participant met exclusionary criteria and, thus, all participants providing consent participated in the study. Demographic information including age, race/ethnicity, gender, native language, and special education services and classification was collected (see Appendix D).

All students in the classrooms (n = 119) participated in the administration of the Forchelli Following Directions Task (FFDT) to assess reliability. This was considered a classroom-based activity, where no demographic or identifying information were collected on these participants; therefore, consent was not deemed necessary. A total of 70 parents consented and students assented to participate in the individual test administration portion of the study. One student was excluded from analysis due to inability to complete tasks. This participant had a documented developmental disability and had difficulty comprehending and following any verbal directions, suggesting his/her performance would be not representative of his/her true ability. Therefore, a total of 69 participants were analyzed. This exceeded the necessary amount of participants required from an *a priori* power analysis conducted; a sample size of around 60 participants was required for a medium effect size (d = .6).

The sample was 58% male (n = 40) and 42% female (n = 29). The average age of a participant was 6 years old (*M* = 6.66, *SD* = 1.09). Table 3.2 provides the means and standard deviations of participants' age by grade. The identified ethnic breakdown of the sample was 42% Caucasian-American/White, 37.7% African American/Black, 5.8% Hispanic, 7.2% Asian American, and 7.2% identified as "Other." Table 3.3 provides a visual comparison of the sample to the United States Census Bureau's national and regional percentages for different ethnic groups (United States Census Bureau, 2013), When considering that the sampling focused on more urban, lower-income populations, mainly in the Philadelphia area, the breakdown is representative (United States Census Bureau, 2013).

Table 3.2

Mean and standard deviation of participants' age and gender by grade

Grade		Gender (F = Female, M = Male)	Age	
	n	*n*	*M*	*SD*
Kindergarten	23	F = 16 M = 7	5.65	.57
1st Grade	29	F = 15 M = 14	6.55	.51
2nd Grade	15	F = 9 M = 6	8.00	.01
3rd Grade	2	F = 0 M = 2	9.50	.71

Table 3.3

Percentages of participants' ethnicities within sample compared to national and regional United Census Bureau

Race/Ethnicity	Current Sample	Philadelphia Regional U.S. Census	National U.S. Census
Caucasian-American/White	42%	45.5%	77.7%
African-American/Black	37.7%	44.2%	13.2%
Latino-American/Hispanic	5.8%	2.4%	17.1%
Asian-American	7.2%	6.9%	5.3%

Measures

Screening/Demographic Information

Parents completed a background questionnaire in English seeking information about child's age, race/ethnicity, gender, native language, and special education services and classification (see Appendix D). This was mailed home to parents with the consent form.

Working Memory

Behavioral Regulation Index of Executive Functions (BRIEF) - Teacher and Parent forms (Gioia, Isquith, Guy, & Kenworthy, 2009). The BRIEF is an 86-item behavior checklist that measures various abilities considered within the executive functions construct. Raters assess the child's behavior on a 3-point Likert-like scale. Raw totals are obtained and converted into T-scores for Emotional Control, Initiate, Inhibit, Organization of Materials, Plan/Organize, Self-Monitor, Shift, and Working Memory. Factor scores are also provided for the Metacognition and Behavioral Regulation indexes, as well as a Global Executive Composite (see Appendix G). The manual indicates that higher T-scores (i.e., above a score of 60) indicate more difficulty within an area. Sample questions from each section include:

Inhibit: "Needs to be told 'no' or 'stop that'"

Shift: "Cannot get a disappointment, scolding, or insult off his/her mind"

Emotional Control: "Overreacts to small problems"

Initiate: "Is not a self-starter"

Working Memory: "When given three things to do, remembers only the first or last"

Plan/Organize: "Does not bring home homework, assignment sheets, materials, etc."

Organization of Materials: "Loses lunch box, lunch money, permission slips, homework, etc."

Monitor: "Does not check work for mistakes"

Please note that it was decided that the full BRIEF would be administered to both teachers and parents to increase validity. Since the measure and standardized scores were normed through participants' completion of the full measure, I wanted to mirror this procedure. Further, I wanted to compare the screener to all areas of executive functions included in this task.

Digit Span Forward (DSF) and Backwards (DSB) (Weschler, 2003). Digit Span is an individually administered objective measure of verbal WM. It has high internal consistency reliability (Cronbach's alpha = .87). In the forward condition of this measure, participants hear a sequence of digits at a rate of one digit per second and are asked to immediately repeat the sequence in the exact order it was presented. The length of the sequence begins with two digits and becomes increasingly more difficult (up to nine digits in length) until the participant receives two errors in a set. In the backwards condition, participants hear a sequence of digits at a rate of one digit per second and are asked to repeat the sequence in the reverse order that is was presented. Similar to DSF, DSB begins with two digits in length and increases to nine digits in length (see Appendix H).

Forwards and Backwards Spatial Span (Weschler, 2003). Spatial Span is an individually administered subtest from the WISC-IV integrated that is thought to tap an individual's visual-spatial WM. The measure utilizes a board with randomly located blue blocks attached to it. In the forward condition, the individual touches a increasingly difficult sequence of blocks in the same order that was modeled by the examiner. In the backward condition, the individual is required to reverse the order of the block sequence modeled by the examiner. For each condition, the length of the sequence begins with two blocks and becomes increasingly more difficult (up to eight blocks in length) until the participant receives two errors in a set.

Forchelli Following Direction Task (FFDT). A group-administered 20-item following directions task was developed by adapting or creating questions that are similar to the following direction tasks utilized in the study by St. Clair-Thompson and colleagues (2010), Gathercole and colleagues (2008), and other standardized direction following tasks (e.g. Understanding Directions from the Woodcock Johnson Tests of Cognitive Abilities - Third Edition (2003), Comprehension of Instructions from the NEPSY-II (Korkman, Kirk, & Kemp, 2007), and the Token Test for Children (McGhee, Erler, & DiSimoni, 2007)). Similar to the task developed by St. Clair-Thompson and colleagues (2010), the items will systemically increase in difficulty, according to number of behaviors and qualifiers. For example, "point to the picture on top of page 3 and copy it twice" is indicated to have two behaviors (i.e., "point" and "copy") and three qualifiers ("on top," "page 3," and "twice"). The initial version was reviewed by three elementary school teachers to assess verisimilitude and verdicality (see Appendix I). Items were also reviewed by researchers with a strong background and familiarity in WM research to assess appropriateness (see Appendix J).

Once amendments were made, the classroom teacher administered the latest version as the researcher observed. All directions were played via an audio recording to ensure fidelity across teachers and classrooms. Once the recording began, the administrator would pause the recording for five seconds between each item number to allow students time to respond. Raw scores are obtained by counting the number of correct responses to the items presented; partial credit is given for correctly responding to part of the item (e.g., only writing a square after the item requested the individual draw a square and a circle).

Procedure

Development/Pilot Study

Test content or content validity was analyzed through communication with experts in both the ecological and behavioral aspects of WM. Three elementary school teachers were asked to review the instructions and items of the FFDT to assess for ecological relevance and provide feedback on a brief questionnaire (Appendix I). Generally, teachers viewed the content of items as appropriate for elementary school-aged children. Teachers provided feedback to change directions on certain items to have more age-appropriate language. More specifically, the teachers requested changing more difficult language. One teacher reported "words like 'third' and 'fifth' may be confusing." Another stated other prepositions, such as "within" to "drawing a line *on* something" were difficult. The respondents provided suggestions, such as simplifying tasks to "put an 'x' or circle something." Teachers also reported that questions toward the end of the task were "complex" and "difficult for younger children."

After edits were made based on the teacher recommendations, the second draft of FFDT was reviewed by four different researchers with great familiarity in the research and assessment of working memory. Three of the reviewers were doctoral-level faculty within the College of Education at Temple University. The fourth reviewer was an independent researcher from an unaffiliated university to control for some bias in response. These faculty provided feedback on the content validity of the measure utilizing the feedback form created (Appendix J). There were minimal changes requested, namely in the language used to promote uniform scoring of items. In addition, experts expressed concern about the other processes required to complete the tasks which may confound results. These included auditory processing, language development, and motor development. These concerns will be addressed further in the discussion chapter of this paper.

Finally, the third version was piloted on two elementary school children by their classroom teachers to assess any final difficulties surrounding administration and instruction clarity. One kindergarten student and one third grade student were utilized to assess difficulty. The students completed the FFDT in less than ten minutes. Anecdotal feedback suggested that the task was very difficult for the kindergarten student and less difficult for the third grade student. Teachers and students did not express any further concerns regarding language or administration.

A brief pilot study evaluated the acceptability of items initially developed for the following directions measure. A group of two elementary school student volunteers were administered the original 15-item following directions measure. After administration, the researcher spoke to participants to gauge the clarity and comprehension of items. Edits were then made based on feedback.

Preliminary Validity Study

During recruitment time, where participants' parents signed and returned the consent forms and completed the demographic questionnaire and BRIEF rating scale, the teachers from the participating schools were provided a brief overview on the group administration of the following directions measure. This involved a five to ten minute meeting reviewing how to administer the audio recording with consideration for five-second response times between items. In addition, an emphasis was made on not responding or aiding students in their responses to items.

After the consents were collected, the researcher scheduled a day with the teacher to administer the group and individually-administered tasks. Initially, all classroom students were given the group administered task. Each student was supplied with a response sheet and the teacher began the audio recording of the directions and test items (see Appendix E). As specified previously, the teacher allowed five seconds for participants to respond to each item. After the last item was read, the teacher collected the response sheets and returned them to the investigator for scoring and analysis. Each student was provided a reinforcement (i.e., small sticker) for their participation. Afterwards, each participating student was brought to a quiet and well-lit room for the administration of the individual WM assessment, by way of the WISC-IV Digit Span and Spatial Span tasks by the researcher. Participants were provided a tangible reinforcement (i.e., small sticker) after completion of each task. The teacher also completed a BRIEF Teacher Rating Form for each participating student. At the completion of scoring all students' data, the teacher and parent of participating students were provided with a score sheet providing an overview of their performance on the standardized measures utilized (Appendix F).

Hypotheses

Overall, the researcher hypothesized that the FFDT would be reliable and valid measure of WM. The reliability of the FFDT would equal or surpass a Cronbach's alpha of 0.70 in this study. A large positive correlation between Digit Span and Spatial Span scores and the FFDT Total Score was hypothesized, as previous research supports a strong relationship between ecologically valid and direct measures of WM (Forchelli, Cariss, & Fiorello, unpublished manuscript). Smaller, negative (because higher scores on the BRIEF indicate more difficulty) correlations were hypothesized to be found between the FFDT and the BRIEF Working Memory Index standard scores due to evidence suggesting that behavioral measures of internal cognitive states are not as representative of ability (Kolko & Kazdin, 1993; Normand & Tannock, 2012). It was also hypothesized that the FFDT would not correlate with other BRIEF subscales, as they measure different constructs.

The relationship between the performance on the FFDT was graphed according to age to assess the relationship to normal WM development. As was discussed previously, WM appears to gradually increase throughout development (Best & Miller, 2010; Garon, Bryson, & Smith, 2008; Marsh, Gerber, & Peterson, 2008). Therefore, it was hypothesized that there would be a small, positive Pearson correlation between the FFDT total score and a participants' age.

CHAPTER FOUR

RESULTS

Preliminary Validity Study

Internal structure or internal reliability of the FFDT was assessed through the use of Cronbach's alpha coefficient (Cronbach, 1988). A higher alpha indicates a higher rate of inter-correlations between test items, indicating higher internal consistency or reliability. This, in conjunction of other evidence, supports that the test is measuring a unitary construct, which in this case is WM. Bracken (1987) outlined guidelines for acceptable alpha coefficient ranges, according to the type of test being investigated. For a group-administered task utilized to make programmatic decisions, Bracken outlines that an alpha level at 0.60 or greater has sufficient internal consistency for a group-administered task; the level increases to .8 for diagnostic screeners administered on an individual basis. More recent research has cited that an alpha level between .70 and .80 is appropriate for educational decisions (Salvia & Ysseldyke, 2004).

Initially, reliability was run on all fifteen items of the FFDT scale ($\alpha = .659$). Upon closer inspection, the first question on the FFDT had an item-total correlation of $p < .001$, indicating no contribution of variance. This question was removed in all subsequent analyses. The second analysis on the 14-item scale produced a Cronbach Alpha of .790. According to developed standards, the FFDT is within the range of good reliability. Inter-rater reliability was also conducted with 20% of the FFDT reliability sample. There was 100% consistency between raters.

The total raw score from the FFDT was compared, using Pearson correlations, to Forwards and Backwards Digit Span, Forwards and Backwards Spatial Span, and the

BRIEF Parent and Teacher Rating Scale Subscale Index T-scores. Table 4.1, Table 4.2 and Table 4.3 provide the correlations between these variables.

Table 4.1

Pearson correlations between the FFDT and raw scores on performance measures

Performance Measure	FFDT Total Score
Forwards Digit Span	.657**
Backwards Digit Span	.629**
Forwards Spatial Span	.658**
Backwards Spatial Span	.639**

$^*p<.05$ $^{**}p<.01$ $^{***}p<.001$

Table 4.2

Pearson correlations between the FFDT and Parent ratings on the BRIEF Rating Scale T-Scores

	FFDT Total Score
Inhibition	-.121
Shift	-.003
Emotional Control	-.034
Initiate	-.034
Working Memory	-.249*
Planning	-.205
Organization	-.052

Table 4.2, continued

Monitor	-.013

*p<.05 **p<.01 ***p< .001

Table 4.3

Pearson correlations between the FFDT and Teacher ratings on the BRIEF Rating Scale T-Scores

	FFDT Total Score
Inhibition	-.163
Shift	.166
Emotional Control	.143
Initiate	-.236
Working Memory	-.247*
Planning	-.225
Organization	-.073
Monitor	-.146

*p<.05 **p<.01 ***p< .001

As indicated in Table 4.1, scores on the FFDT were significantly and positively correlated with performance measures of WM, including digit span and spatial span tasks. That is, participants performed similarly on Digit Span, Spatial Span, and the FFDT. As hypothesized, Table 4.2 & Table 4.3 show weaker negative correlations with parent and teacher ratings of working memory with the FFDT. This indicates that children who performed worse on the FFDT also had higher clinical indications of difficulties with WM by parents and teachers. The FFDT was not correlated with any other executive function subscales on parent and teacher rating scales. This suggests that the FFDT score was specific to WM difficulty and remained unrelated to performance in other areas of executive functions.

The relationship between age and the FFDT was explored. A strong positive correlation was found between the total FFDT and participants' age in months ($r(67) = .66$, $p<.001$). This indicates that as students' age increased, their performance on FFDT increased (Figure 4.1).

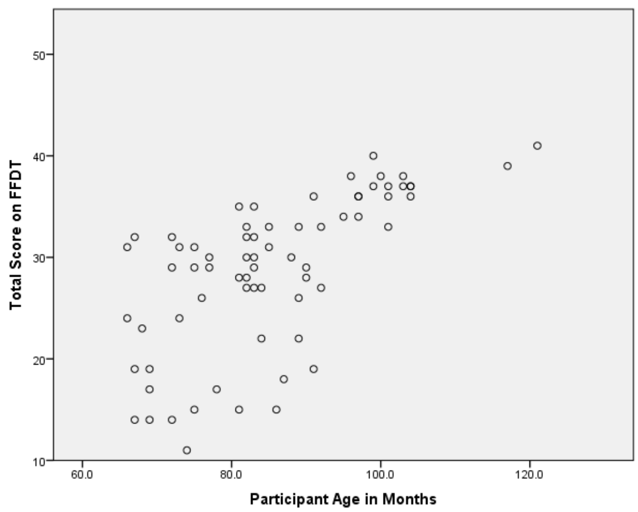

Figure 4.1. Scatterplot of FFDT Total Score by Participants' age in months

To further investigate the relationship between development and FFDT performance, a one-way ANOVA was performed between the FFDT total score and students' grade. This analysis only included kindergarten, first, and second grade students due to the low number of participants in third grade (n = 2). There was a significant interaction between grade and performance on the FFDT ($F(3, 65) = 22.05$, $p<.001$, $\eta^2 = .51$). Table 4.4 shows the Tukey Post Hoc on this analysis.

Table 4.4

Tukey post hoc analysis mean difference and standard deviation between participant grade level and FFDT total score

	Kindergarten	1st Grade
1st Grade	-5.72 (1.53)**	
2nd Grade	-13.80 (1.82) ***	-8.1 (1.74)***

$*p<.05$ $**p<.01$ $***p< .001$

The Tukey results demonstrate that participants in higher grade levels performed significantly better than participants in lower grades. That is, participants in kindergarten scored significantly lower (*m* = 22.87, *sd* = 7.2) than participants in first grade (*m* = 28.59, *sd* = 5.25) and second grade (*m* = 36.67, *sd* = 1.68).

A Receiver Operator Characteristic (ROC) Curve analysis was used to examine the diagnostic accuracy of the FFDT in identifying participants with WM difficulty similar to performance-based measures. This analysis is used to determine the sensitivity and specificity of a new measure against existing diagnostic tools. In this case, the FFDT will be compared against the performance-based measures, Digit Span Backwards and Spatial Span Backwards, in identifying WM difficulty. WM difficulty on the performance-based measures used was defined as a scaled score below an eight, representing performance below the 25^{th} percentile. This cutoff is used in similar screening tools, such as Curriculum Based Measures (Shinn, 1998); it demonstrates a performance below 75% of the sample. This indicates these participants were in the lowest quartile of performance. This cutoff was, then, re-coded into a binary, with "0" identified as not meeting criteria for WM difficulty (i.e., scaled score greater than 7) and

"1" identified as meeting criteria for WM difficulty (i.e., scaled score less than or equal to 7). Table 4.5 provides the statistics associated with this analysis.

Table 4.5

Diagnostic Utility Statistics for the FFDT

Variable	Overall *p* Values	AUC	Score Comparisons
Digit Span Backwards Scaled Score	.003	.76	
Spatial Span Backwards Scaled Score	.042	.66	1<8

Note: AUC = Area Under the Curve.

The Area Under the Curve (AUC) indicates that the FFDT has a medium effect size in determining WM difficulty on the Digit Span Backwards task (Rice & Harris, 1995). This demonstrates that 76% of the time, students with WM difficulties performed more poorly on the FFDT than students without WM difficulties. A low effect size was found between FFDT and the Spatial Span Backwards task. This demonstrates that 66% of the time, students with WM difficulties performed more poorly on the FFDT than students without WM difficulties. Figure 4.2 and Figure 4.3 present a visual representation of these data.

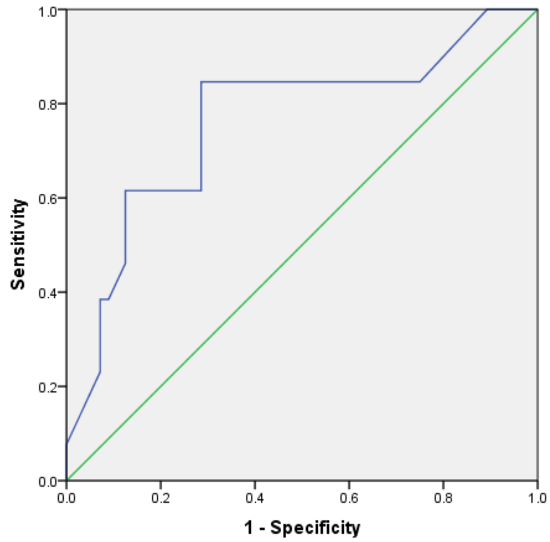

Figure 4.2. Receiver Operating Characteristic Curves for Scaled Score below 25th percentile on Digit Span Backwards

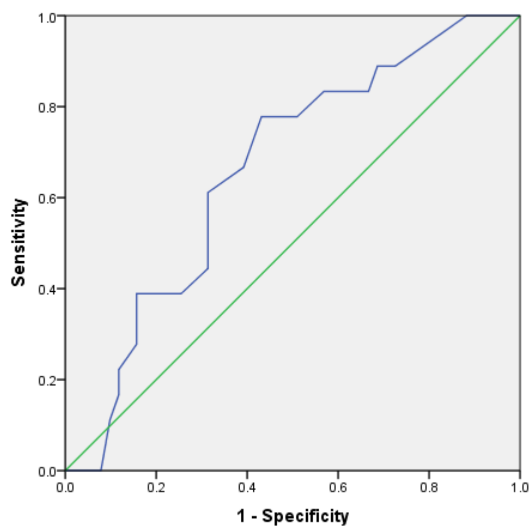

Figure 4.3. Receiver Operating Characteristic Curves for Scaled Score below 25th percentile on Spatial Span Backwards

A curve farther away from the diagonal trend line indicates a greater AUC and greater sensitivity and specificity of the measure. Both Figure 4.2 and 4.3 demonstrate a low to medium AUC, indicating that the FFDT has specificity and sensitivity to identifying WM difficulty compared to existing diagnostic tools.

CHAPTER FIVE

DISCUSSION

Overall, results provide preliminary evidence of the validity of FFDT scores for screening for WM difficulties in elementary-aged children. This study provides validity evidence across four areas outlined by AERA, APA, and NCME (2014), namely internal structure, test content, response process validity, and construct validity. Further, ecological validity is suggested.

First, the FFDT demonstrated strong internal consistency. This suggests that the developed task measures a unitary construct. The data collected from the pilot and validity studies support that WM is the measured construct within the FFDT. Four independent reviewers indicated content validity. They agreed that to receive credit on items of the FFDT task, respondents must utilize their WM and exert increased effort as items increase. There is some potential for bias with three of the four expert reviewers, as they have a professional relationship with the researcher. Social desirability may have influenced the views of these reviewers and caused them to indicate more favorable views of the measure. However, the fourth responder was taken from an independent institution and had no personal relationship with the researcher. This reviewer's responses were similar to the other researchers, suggesting that the overall evaluation of the FFDT was accurate.

Teacher and student feedback in the pilot study suggested that the FFDT task demonstrates response process validity, as well as ecological validity. Teachers agreed that the items involved directions and tasks that elementary school students execute in the classroom. While teachers did provide suggestions to make language more appropriate on the items, they agreed that globally, the directions asked on the FFDT reflect activities within their classrooms. They also indicated that items became increasingly difficult for younger students, which was intended to lessen chances for a ceiling effect in older or more advanced students. When considering these responses, it can be argued that the FFDT mirrors the utilization of WM within the school setting and supports ecological validity of the measure, as is urged by Levin, Thurman, and Kiepert (2010). Further, the two pilot participants' performance provided preliminary evidence that the task held response process validity and external validity. The students completed the task in less than ten minutes, making it both practical and feasible for use within the classroom.

The comparative analyses provide support for construct validity. Strong positive relationships were found between the FFDT and existing performance measures of WM, namely Digit Span and Spatial Span. Traditionally, Digit Span is viewed as assessment of verbal WM and Spatial Span as an assessment of visual/spatial WM. Results indicated similarly strong relationships between the FFDT with both measures, suggesting relationships with both auditory and visual/spatial WM. This suggests that FFDT assesses both areas of WM as well as existing measures.

Further, positive relationships were found between FFDT and the WM indexes on both the parent and teacher behavioral rating scales. As discussed previously, researchers question the accuracy of rating scales in identification of WM difficulties. It can be difficult for outside respondents to observe and ascertain the level of proficiency of internal states and cognition. This is discussed in past research by Kolko and Kazdin (1993). Normand and Tannock (2012) specifically criticize the main rating scale for executive function assessment, the BRIEF (Gioia, Isquith, Guy, & Kenworthy, 2000) for not adequately pinpointing difficulties with WM. Notwithstanding, this research demonstrates a small but statistically significant relationship with teacher and parent behavioral observations of WM skill and performance on the FFDT. As more difficulties are observed in WM, there is poorer performance on the FFDT. This suggests that the BRIEF rating scale has some utilization in WM difficulty detection.

No relationships were found with other executive function indexes on the BRIEF. There is a strong relationship between WM and other EF skills, as highlighted in research by St. Clair-Thompson (2011), which creates the possibility for other elevations on the BRIEF. The lack of significant relationships between other EF subscales on the BRIEF, however, indicates specificity in the FFDT to assess WM ability over other executive functions. This provides further support that not only the FFDT scores can be validly used to screen for WM difficulties, it has specificity to only identify children with WM difficulties. This is supported by the ROC analysis indicating low to medium power in the FFDT positively identifying WM difficulty in the sample.

The FFDT appears to have some sensitivity to a child's developmental level. That is, performance increased as age and grade increased. This mirrors past research demonstrating gradual increases in WM ability (Brocki & Bohli, 2004; Gathercole et al., 2004; Wu et al., 2011). Notwithstanding, a limitation of this study was the smaller number of older elementary students, indicated with the sample's mean age around six years old. This may indicate that the results from this study are only generalizable to certain aged students, since the majority of participants in this study were kindergarten and first grade students. In order to investigate this relationship further, future studies should acquire a larger sample of second and third grade students (i.e., ages eight through ten years old). However, as this is intended to be an early screener, usefulness in Kindergarten and first grade is valuable, despite the weaker evidence for use with older students.

The expert panel voiced some concern on the confounding variables that appear inherent in the task. In particular, they questioned the inability to differentiate if poor performance was an indicator of true WM difficulty or other difficulties involved in WM, mainly language and listening comprehension. As discussed in earlier chapters, language and WM are closely related. Baddeley (1996) argued that memory difficulties appear underlying to early language acquisition, and lower WM may be underlying the language-based difficulties (Archibald & Gathercole, 2006; Leonard et al., 2007). This close relationship makes it difficult to ascertain in any verbal task which is truly affecting performance. However, research suggests that more than likely, poor performance is related to WM, as students with language difficulties also inherently have poor WM.

Listening comprehension is more related to an individual's ability to process information and relate it to pre-existing knowledge bases and educational experiences (Flanagan, Ortiz, & Alfonso, 2013); it is considered an academic skill. Crystallized knowledge, a cognitive ability, is related to a student's past learning experiences and impacts listening comprehension. Poor crystallized knowledge may cause false positives on the WM scale. However, if their performance was mainly due to lack of education and knowledge, it can be argued that their performance on the verbal and spatial WM tasks would not have been affected. More than likely, listening comprehension and poor crystallized knowledge highlight the close link between major areas of academic achievement and WM (Alloway et al., 2009). A child with poor WM will more than likely also have poorer crystallized knowledge due to the "bottleneck hypothesis" postulated by Alloway and Gathercole (2006). These individuals encode and retain less information into their LTM during instruction, which weakens their knowledge base and impacts future performance.

Overall, it should be highlighted that the intention of the FFDT is as a screener and not a diagnostic indicator of WM difficulty. In order to adequately assess WM ability within the classroom with mindfulness to social and ecological validity, verbal directions were utilized, similar to studies conducted by Gathercole and colleagues (2008). Development of the FFDT involved teacher input to identify directions that appeared to demonstrate verisimilitude and were developmentally appropriate for general education students. This indicates that while there is some loading on language in the task, these items should not measure language, alone. If poor performance is indicated, the child is demonstrating difficulty that is not normative for the general student population and would indicate a need for further assessment.

In this study, the FFDT demonstrated correlations with both verbal and visual WM. If language was the main area of measurement, there should not be a strong correlation between both; language difficulty should differentially affect performance on the WM task and not the visual WM task. This suggests that WM was more likely the construct being measured than language difficulties alone.

The purpose of the FFDT is to flag or raise concern for students that demonstrate the difficulty to hold and manipulate information presented (WM). This is due to its extraordinary importance for a developing learner. Most recent research highlight WM as a major component to "cognitive efficiency" as a mediator to all cognitive skills (Taub

& McGrew, 2014). It has demonstrated strong links to basic reading, reading comprehension, basic math, and math reasoning skills across all age ranges (Flanagan et al., 2006; McGrew & Wendling, 2010). Due to its importance across academic skills, McGrew and Wendling (2010) recommend utilizing it as an earlier screener to identify "precursors" to more serious difficulties in the future. Notwithstanding, similar to diagnosing other cognitive and neuropsychological difficulties, an individual should receive more comprehensive assessment to ascertain and parse out the possibility of other related disorders or difficulties.

Limitations of the Study

There are a number of limitations to this study that should be highlighted. One of the largest limitations is related to the sample used. This study was based on a convenience sample, where schools were selected mainly on the administration's willingness to participate. The majority of participating schools were privately run and had involved principals and teachers who desired involvement in research. This selection bias has potential to skew results. Further, the use of mainly private institutions is problematic. Private institutions suggest higher level of social support of students due to the need to apply and financially fund attendance at these institutions. More attentive and supportive school environments have been found to influence student performance (Aiens & Barbarin, 2008). These schools may have more resources to adequately support the academic needs and differential learning of each student. Conversely, schools with less administrative and teacher involvement may house students with exposure to less evidence-based instruction and support. In the latter profile of school, students may perform significantly differently on the FFDT and demonstrate that the measure has less specificity or sensitivity in these populations.

Further, within each school unit, only one or two classrooms were used. It cannot be ascertained that performance in each of these classrooms is representative of the entire school. This sample of students, then, may not represent students within an urban public school district or a more rural district with less social and financial supports. This limits generalization of findings to the broader urban area of Philadelphia or to a larger regional or national level. It is hoped that future studies will have more incentives to attract a more inclusive sample of schools.

As discussed previously, there also is more representation of Kindergarten and first grade students than in later grades. This may impact the representation of FFDT on performance in second or third grade. The few third grade students utilized received almost perfect scores on the measure, which may suggest a ceiling effect in higher grades and inability to assess the range of skills in students in that age group. Further, the sample utilized does appear more representative of the ethnic breakdown of the Philadelphia metropolitan area, but not a national sample. While this was the intention of this study, it may not generalize at the state or national level.

There also may be concerns in the validity across administrations. While the researcher observed the administration by teachers, there was no formal evaluation of administration fidelity. Therefore, there might have been inherent differences in how teachers supported the students in each classroom which may have influence responding. Demographic information on teachers and parents were also not collected and, therefore, no formal evaluation of administration fidelity. This may make which makes replication of this study difficult.

Further, there is the potential for examiner bias. This researcher was responsible for all data collection and scoring with limited involvement of outside evaluation. It can be argued that, during scoring procedures, the examiner unconsciously viewed students less favorably to promote a larger percentage of WM difficulties within the sample. It is hoped that this was somewhat controlled by standardization of scoring procedures on both the existing performance measures, as well as scoring procedures developed prior to data collection. The inter-rater reliability does provide some support for lack of scoring bias; however, only future independent research will fully validate this. Further, teachers and parents of participants were provided some information on what the FFDT measure was investigating which may have influenced how they responded to the rating scale provided. This may have impacted how they rated severity of symptoms related to the WM scale and possibly minimized other behavioral difficulties related to other EF skills.

When considering the development of the measure, there are still areas of validity that should be assessed. The FFDT needs to show convergent and divergent validity to related variables. As in the case of auditory processing and crystallized knowledge, the FFDT should show correlation to some abilities but also show enough differential performance that they are demonstrated as different constructs. This can be achieved through comparing performance on the FFDT to existing measures of cognitive abilities and executive functions across age groups.

Applications to Practice

The FFDT provides a new tool for both educators and psychologists in supporting students. First, it is more time efficient than other performance-based measures. During this study, the FFDT took up to fifteen minutes for a teacher to complete on a classroom of twenty to thirty students. Conversely, the individually-administered subtests took approximately fifteen minutes *per* child to complete by a graduate-level clinician, equating to approximately five hours to assess a classroom of twenty students. The latter removes the child from classroom activities for the time of assessment and requires school psychologist involvement in each student's assessment. This suggests that the FFDT takes less time away from classroom instruction for the teacher, as well as less time for the school psychologist in preliminary assessment of students.

The recent trend in education to progress monitoring also supports the use of FFDT. The Response to Intervention (RtI) framework supports the baseline assessment and monitoring of improvement (or lack thereof) of students throughout their educational careers. Presently, there are no measures that are used to monitor growth of executive function skills, however, they are critical for learning in students. As stated previously, McGrew and Wendling (2010) suggest utilizing a screener such as the FFDT to identify "precursors" to more broad-band difficulties in learning. WM screening would not only benefit for identification; it also may guide remediation. Research has demonstrated some support that WM can be trained or improved with proper remediation (Holmes, Gathercole, & Dunning, 2009). Computer programs such as Cogmed (Klingberg et al., 2005) have demonstrated increasing popularity and usage in schools. Therefore, the FFDT may be a tool to acquire baseline data and monitor progress in a tiered intervention plan for students. This will also flag students for teachers to focus on or monitor more closely.

While this may be beneficial to provide quick and accurate assessment of WM, once screened, it will be important to understand the impact of other variables involved in WM development and training. A study by Jaeggi, Buschkuehl, Shah and Jonides (2014) demonstrated that variables such as motivation impact the success of training. The interplay of all these aspects should be researched more thoroughly.

Further, in creating a screener only able to detect WM ability, it does not cover the broad scope of other variables that may be important to screen and monitor in early childhood. The relationship between the FFDT and EF skill should be further explored in future studies. Comparisons should be initiated between the FFDT performance and performance-based measures of other executive function skills.

Future Research

Further information should be collected to continue to support the reliability of the FFDT measure. Temporal stability of the measure can be calculated through test-retest reliability. This will provide information regarding the consistency of performance on the FFDT over time or over multiple administrations. Temporal stability will be particularly important to demonstrate in the FFDT for usage in an RtI framework. The FFDT must demonstrate the ability to adequately test WM ability over time and demonstrate sensitivity to change over multiple administrations. In addition, it may be beneficial to develop an alternative form of the FFDT to allow better temporal stability. This will help control for testing effects on performance.

As suggested in previous sections, the FFDT also needs to be compared to performance on other measures to assure convergent and divergent validity. Due to concerns by experts in this study, the FFDT should be compared to performance in cognitive and academic skills. It can be achieved by comparing the FFDT performance to standardized measures of intelligence and academic achievement. This will be important to provide support for its construct validity. Further, it may be interesting to have a better understanding of its relationship to academic performance measures over time. The FFDT should also be compared to other performance-based EF measures to ascertain its relationship to other EF constructs.

To control for limitations discussed surrounding method and sample, future studies should look to demonstrate FFDT's validity in larger and more representative populations. If possible, research should look to more appropriate sampling methods, such as utilizing an entire school population or sampling across all grades in a school district. It is hoped that with time, the FFDT can also be sampled across varying geographic locations and distributions of population (i.e., urban, suburban, or rural). It also should look to have equal representation of differing age groups, particularly of older students, to confirm improved performance with age.

Due to WM's relationship to a variety of clinical populations, such as those with specific learning disabilities or ADHD, performance on the FFDT should also be compared in these populations. This may provide interesting implications of diagnostic considerations in these subgroups. These studies may be accomplished with the use of monetary incentives for school districts or the development of research relationships with varying universities to distribute labor/access.

REFERENCES

Adams, A., & Gathercole, S. E. (1995). Phonological working memory and speech productino in preschool children. *Journal of Speech & Hearing Researcher, 38(2)*, 403.

Aikens, N. L., & Barbarin, O. (2008). Socioeconomic differences in reading trajectories: The contribution of family, neighborhood, and school contexts. *Journal of Educational Psychology, 100***,** 235-251.

Alloway, T. P. (2011). A comparison of working memory profiles in children with ADHD and DCD. *Child Neuropsychology, 17(5)*, 483-494.

Alloway, T. P., Gathercole, S. E., Kirkwood, H., & Elliott, J. (2009). The working memory rating scale: A classroom-based behavioral assessment of working memory. *Learning and Individual Differences, 19,* 242–245.

Alloway, T. P., Gathercole, S. E., Kirkwood, H., & Elliott, J. (2009). The cognitive and behavioral characteristics of children with low working memory. *Child Development, 80(2),* 606-621.

Alloway, T. P., Gathercole, S. E., Willis, C., & Adams, A. (2004). A structural analysis of working memory and related cognitive skills in young children. *Journal of Experimental Child Psychology, 87,* 85-106.

Alloway, T., Gathercole, S. E., & Kirkwood, H. (2008). Working Memory Rating Scale.

Alloway, T., Gathercole, S. E., Kirkwood, H., & Elliot, J. (2008). Evaluating the validity of the automated working memory assessment. *Educational Psychology: An International Journal of Experimental Educational Psychology*, *28(7)*, 725-734.

Alloway, T.P. & Gathercole, S.E. (2006). How does working memory work in the classroom. *Educational Research and Review, 1(4)*, 134-139.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (2014). *Standards for Educational and Psychological Testing*. Washington, DC: American Educational Research Association.

Archibald, L. M. D., & Gathercole, S. E. (2006). Nonword repetition in specific language impairment. *Psychonomic Bulletin and Review, 14*, 919-924.

Baddeley AD, Allen RJ, Vargha-Khadem F. 2010. Is the hippocampus necessary for visual and verbal binding in working memory? *Neuropsychologia, 48*, 1089–95.

Baddeley, A. D. (1996). Exploring the central executive. *Quarterly Journal of Experimental Psychology, 49A(1)*, 5–28.

Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Science*, *4*(11), 417–423.

Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology, 63, 1-29.*

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

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

Baddeley, A. D., Allen, R. J, & Hitch, G. J. (2011). Binding in visual working memory: The role of the episodic buffer. *Neuropsychologia, 49*, 1393–400.

Baddeley, A. D., Grant, W., Wight, E., & Thomson, N. (1975). Imagery and visual working memory. In P.M.A. Rabbitt, & S. Dornic (Eds.) *Attention and Performance Visualization* (pp. 205–17). London: Academic.

Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions. *Psychological Bulletin, 121*, 65-94.

Barkley, R. A., Edwards, G., Laneri, M., & Fletcher, K. (2001). Executive functioning, temporal discounting, and sense of time in adolescents with attention deficit hyperactivity disorder (ADHD) and oppositional defiant disorder (ODD). *Journal of Abnormal Child Psychology, 29*, 541-556.

Baron, I. S. (2004). *Neuropsychological evaluation of the child*. New York, NY: Oxford University Press.

Beck, D. M. (2010). The appeal of the brain in popular press. *Perspectives on Psychological Science, 5(6), 762-766.*

Berninger, V. W., Abbott, R. D., Swanson, H. L., Lovitt, D., Trivedi, P., Lin, S. C., Gould, L., Youngstrom, M., Shimada, S., & Amtmann, D. (2010). Relationship of word- and sentence- working memory to reading writing in second, fourth, and sixth grade. *Language, Speech, and Hearing in School, 41*, 179-193.

Best, J. R., & Miller, P. H. (2010). A developmental pespective on executive function. *Child Development, 81(6)*, 1641-1660.

Botting, N., & Conti-Ramsden,, G. (2001). Non-word repetition and language development in children with specific language impairment. *International Journal of Language and Communication Disorders, 36(4)*, 421-432.

Boudreau, D., & Costanza-Smith, A. (2011). Assessment and treatment of working memory deficits in school-age children: The role of the speech-language pathologist. *Language, Speech, & Hearing Services in Schools, 42*, 152-166.

Broadbent, D. E. (1958). *Perception & communication.* London: Pergamon Press.

Brocki, K. C., & Bohlin, G. (2004). Executive functions in children aged 6 to 13: A dimensional and developmental study. *Developmental Neuropsychology, 26(2),* 571-593.

Brocki, K. C., Eninger, L., Thorell, L. B., & Bohlin, G. (2010). Interrelations between executive function and symptoms of hyperactivity/impulsivity and inattention in preschoolers: A two year longitudinal study. *Journal of Abnormal Child Psychology*, *38*, 163-171.

Bronfenbrenner (1997). Ecological Models of Human Development. In M. Gauvain & M. Cole (Eds.) *Readings on the Development of Children* (4th ed.)(pp. 37-43), New York, NY: Worth Publishers.

Brown-Chidsey, R. & Steege, M. W. (2005). *Response to intervention: Principles and strategies for effective practice*. New York: Guilford.

Brunswik, E. (1956). *Perception and the representative design of psychological experiments*. Berkeley: The University of California Press.

Buckner, J. C., Mezzacappa, E., & Beardslee, W. R. (2009). Selfregulation and its relations to adaptive functioning in low-income youths. American Journal of Orthopsychiatry, 79, 19–30.

Bull, R., Espy, K. A., & Wiebe, S. A. (2008). Short-term memory, working memory, and executive functioning in preschoolers: Longitudinal predictors of mathematical achievement at age 7 years. *Developmental Neuropsychology, 33*, 205–228.

Case, L. P., Speece, D., Silverman, R., Ritchey, K. D., Schatschneider, C., Cooper, D. H., Montanaro, E., & Jacobs, D. (2010). Validation of a supplemental reading intervention for first-grade children. *Journal of Learning Disabilities, 43(5)* 402-417.

Chein, J. M. & Morrison, A. B. (2010). Expanding the mind's workspace: Training and transfer effects with a complex working memory span task. *Psychonomic Bulletin & Review, 17*, 193-199.

Chrysochoou, E., Bablekou, Z., Tsigilis, N. (2013). Working memory contributions to reading comprehension components in middle childhood children. *The American Journal of Psychology, 124(3)*, 275-289.

Coley, R. J. (2002). *An uneven start: Indicators of inequality in school readiness*. Princeton, NJ: Educational Testing Service.

Compton, D. L., Fuchs, L. S., Fuchs, D., Lambert, W., & Hamlett, C. (2012). The cognitive and academic profiles of reading and mathematics learning disabilities. *Journal of Learning Disabilities, 45(1),* 79-95.

Conklin, H. M., Luciana, M., Hooper, C. J., & Yarger, R. S. (2007). Working memory performance in typically developing children and adolescents: Behavioral evidence of protracted frontal lobe development. *Developmental Neuropsychology, 31*, 103-128.

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information processing system. *Psychological Bulletin, 104,* 163–191.

Cowan, N. (1999). An embedded-processes model of working memory. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 62–101). Cambridge, England: Cambridge University Press.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24,* 87–114.

Cowan, N., Elliott, E. M., Saults, J. S., Morey, C. C., Mattox, S., Hismjatullina, A., & Conway, A. R. A. (2005). On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. *Cognitive Psychology, 51,* 42–100.

Cowan, N., Li, D., Moffitt, A., Becker, T. M., Martin, E. A., Saults, J. S., & Crist, S. E. (2011). A neural region of abstract working memory. *Journal of Cognitive Neuroscience, 23,* 2852–2863.

Cronbach, L. J. (1988). Five perspectives on validation argument. In H. Wainer & H. Braun (Eds.) *Test validity* (pp. 34-35). Hillsdale, NJ: Erlbaum.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*, 281-302.

Dagan, C. P., Braelen, J., Ferrer, E., & Liu, C. (2012). Unitary or non-unitary nature of working memory? Evidence from its relation to general fluid and crystallized intelligence. *Intelligence, 40*, 499-508.

Dahlin, K. I. E. (2010). Effects of working memory training on reading in children with special needs. *Reading and Writing, 24,* 479-491.

De Beni, R., Palladino, P., Pazzaglia, F., & Cornoldi, C. (1998). Increases in intrusion errors and working memory deficit of poor comprehenders. *The Quarterly Journal of Experimental Psychology, 51A(2),* 305-320.

Dell'Acqua, R., & Jolicoeur, P. (2000). Visual encoding of patterns is subject to dual-task interference. *Memory & Cognition, 28(2),* 184-191.

Deutsch, J. A., & Deutsch, D. (1963). Attention: Some theoretical considerations. *Psychological Review, 70*, 80-92.

Diamond, A., & Doar, B. (1989). The performance of human infants on a measure of frontal cortex function, the delayed response task. *Developmental Psychobiology, 22*, 271-294.

Diamond, A., Prevor, M., Callender, G., & Druin, D. (1997). Prefrontal cortex cognitive deficits in children treated early and continuously for PKU. *Monographs of the Society for Research in Child Development, 62*, 1-208.

Dobbs, D. (2005). Fact or phrenology? *Scientific American Mind, 16(1),* 24-31.

Elliot, J.G., Gathercole, S.E., Alloway, T.P., Holmes, J., & Kirkwood, H. (2010). An evaluation of a classroom-based intervention to help overcome working memory difficulties and improve long-term academic achievement. *Journal of Cognitive Education & Psychology, 9(3)*, 227-250.

ElementarySchools.org (2014). Retrieved from www.elementaryschools.org

Ewing- Cobbs, L., Prasad, M., Landry, S. & Kramer, L. (2004) Executive functions following traumatic brain injury in young children: A preliminary analysis. *Developmental Neuropsychology, 26,* 4487-512.

Farmer, E.M.Z., Burns, B.J., Phillips, S.D., Angold, A., and Costello, E.J. (2003). Pathways into and through mental health services for children and adolescents. *Psychiatric Services,* 54, 60-66.

Flanagan, D. P., Ortiz, S. O., & Alfonso, V. (2006). *The achievement test desk reference, 2nd ed.* New York: Wiley.

Flanagan, D. P., Ortiz, S. O., & Alfonso, V. (2007). *Essentials of cross-battery assessment*. Hoboken, NJ: John Wiley & Sons.

Forchelli, Carris, and Fiorello (unpublished manuscript). Ecological measurement of working memory and prediction of academic achievement.

Frazen, M. D., & Wilhelm, K. L. (1996). Conceptual foundations of ecological validity in neurological assessment. In R. J. Sbordone & C. J. Long (Eds.), *Ecological validity of neuropsychological testing* (pp. 91–112). Delray Beach, FL: GR Press/St. Lucie Press.

Garcia-Madruga, J. A., Elosua, M. R., Gil, L., Gomez-Veiga, I., Vila, J. O., Orjales, I., Contreras, A., Rodriguez, R., Melero, M. A., & Duque, G. (2013). Reading comprehension and working memory executive processes: An intervention study in primary school students. *Reading Research Quarterl, 48(2)*, 155-174.

Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in preschoolers: A review using an integrative framework. *Psychological Bulletin, 134(1)*, 31-60.

Gathercole, Pickering, Ambridge, & Wearing (2004). The structure of working memory from 4 to 15 years of age. *Developmental Psychology, 40(2),* 177-189.

Gathercole, S. E., & Alloway, T. (2008). Working memory and classroom learning. In S. K. Thurman & C. A. Fiorello (Eds.), *Applied cognitive research in K-3 classrooms* (pp. 17–40). New York: Routledge.

Gathercole, S. E., Pickering, S. J., Knight, C., & Stegmann, Z. (2004). Working memory skills and educational attainment: Evidence from national curriculum assessments at 7 and 14 years of age. *Applied Cognitive Psychology, 18,* 1–16.

Gathercole, S. E., Service, E., Hitch, G. J., Adams, A., & Martin, A. J. (1999). Phonological short-term memory and vocabulary development: Further evidence on the nature of the relationship. *Applied Cognitive Psychology, 13(1),* 65-77.

Gathercole, S.E., Alloway, T.P., Willis, C.S., & Adams, A.M. (2006). Working memory in children with reading disabilities. *Journal of Experimental Child Psychology, 93*, 265- 281.

Gathercole, S.E., Durling, E., Evans, M., Jeffcock, S., & Stone, S. (2008). Working memory abilities and children's performance in laboratory analogues of classroom activities. *Applied Cognitive Psychology, 22,* 1019-1037.

Gilchrist, A. L., Cowan, N., & Naveh-Benjamin, M. (2009). Investigating the childhood development of working memory using sentences: New evidence for the growth of chunk capacity. *Journal of Experimental Child Psychology, 104,* 252-256.

Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2000). *Behavior Rating Inventory of Executive Functions (BRIEF)*. Odessa, FL: Psychological Assessment Resources.

Hale, J. B., & Fiorello, C. A. (2004). *School neuropsychology: A practitioner's handbook.* New York, NY: The Guilford Press.

Heckman, J. J. (2006) Skill formation and the economics of investing in disadvantaged children. *Science, 312*, 1900-1902.

Holmes, J., Gathercole, S.E., & Dunning, D.L. (2009). Adaptive training leads to sustained enhancement of poor working memory in children. *Developmental Science, 12(4)*, F9-F15.

Hunt, R. R., & Ellis, H. C. (2004). *Fundamentals of Cognitive Psychology* (7th ed.) New York, NY: McGraw-Hill Higher Education.

Jaeggi, S. M., Buschkuehl, M., Shah, P., & Jonides, J. (2014). The role of individual differences in cognitive training and transfer. *Memory and Cognition, 43*, 464 - 48.

Johnson, M. H. (2005) *Developmental cognitive neuroscience* (2nd ed.) Oxford, England: Blackwell Publishers.

Jonides, J., Lewis, R. L., Nee, D. E., Lustig, C. A., Berman, M. G., & Moore, K. S. (2008). The mind and brain of short-term memory. *Annual Review of Psychology, 59*, 193-224.

Kahneman (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice-Hall.

Kershaw, S., & Schatschneider, C. (2012). A latent variable approach to the simple view of reading. *Reading and Writing*, *25*, 433-464.

Klingberg, T., Fernell, E., Olesen, P. J., Johnson, M., Gustafsson, P., Dahlstrom, K., et al. (2005). Computerized training of working memory in children with ADHD- A randomized, controlled trial. *Journal of the American Academy of Child and Adolescent Psychiatry, 44*, 177-186.

Kolko, D. J. & Kazdin, A. E. (1993). Emotional/behavioral problems in clinic and nonclinic children: Correspondence among child, parent, and teacher reports. *Journal of Child Psychology and Psychiatry, 34(6)*, 991–1006.

Korkman, M., Kirk, U., & Kemp, S. (2007). NEPSY—Second Edition (NEPSY-II). San Antonio, TX: Harcourt Assessment

Kwon, H., Reiss, A. L., & Menon, V. (2002). Neural basis of protracted developmental changes in visuo-spatial working memory. *Proc Natl Acad Sci U S A, 99*(20), 13336-13341.

Lambek, R., & Shevlin, M. (2011). Working memory and response inhibition in children and adolescents: Age and organization issues. *Scandinavian Journal of Psychology, 52,* 427-432.

Latzman, R. D., & Markon, K. E. (2009). *The factor structure and age-related factorial invariance of the Delis-Kaplan ExecutiveFunction System (D-KEFS).*

Leffard, S. A., Miller, J. A., Bernstein, J., DeMann, J. J., Mangis, H. A., & McCoy, E. L. B. (2006). Substantive validity of working memory measures in major cognitive functioning test batteries for children. *Applied Neuropsychology, 13,* 230–241.

Leonard, L. B., Weismer, S. E., Miler, C. A., Francis, D. J., Tomblin, J. B., & Kail, R. V. (2007). Speed of processing, working memory, and language impairment in children. *Journal of Speech, Language, & Hearing Research, 50,* 408-428.

Levick, W. R. (2011). Observer rating of memory in children: A review. *Brain Impairment, 11(2)*, 144-151.

Levin, D. S., Thurman, S. K., & Kiepert, M. H. (2010). More than just a memory: The nature and validity of working memory in educational settings. In G. M. Davies & D. B. Wright (Eds.) *Current issues in applied memory research* (pp. 72–95). New York, NY, US: Psychology Press.

Logie, R. H., & Baddeley, A. D. (1987). Cognitive processes in counting. *Journal of Experimental Psychology, 13(2)*, 310-326.

Loosli, S.V., Buschkuehl, M., Perrig, W.J., & Jaeggi, S.M. (2012). Working memory training improves reading processes in typically developing children. *Child Neuropsychology, 18(1)*, 62-78.

Magimairaj, B. M., & Montgomery, J. W. (2012). Children's verbal working memory: Role of processing complexity in predicting spoken sentence comprehension. *Journal of Speech, Language, & Hearing Research, 55*, 669-682.

Mahone, E. M., Martin, R., Kates, W. R., Hay, T., & Horska, A. (2009). Neuroimaging correlates of parent ratings of working memory in typically developing children. *Journal of International Neuropsychological Society, 15*, 31-41.

Marsh, R., Gerber, A. J., & Peterson, B. S. (2008). Neuroimaging studies of normal brain development and their relevance for understanding childhood neuropsychiatric disorders. *Journal American Academic Child Adolescent Psychiatry, 47(11)*, 1233-1251.

Martinussen, R. H., Hayden, D. C., Hogg-Johnson, S., & Tannock, R. (2005). A meta-analysis of working memory impairments in children with attention-deficit/hyperactivity disorder. *Journal of American Academic Child Adolescent Psychiatry, 44(4)*, 377 - 384.

McGhee, R., Erler, D. J., & DiSimoni, F. (2007). *Token Test for Children – Second Edition,* Oceanside, CA: PRO-ED.

McGrew, K. S. & Wendling, B. J. (2010). Cattell – Horn – Carroll cognitive achievement relations: What we have learned from the past 20 years of research. *Psychology in Schools, 47(7),* 651-675.

Messick, S. (1995). Validity of psychological assessment: Validation of interferences from persons' responses and performances as scientific inquiry into score meaning, *American Psychologist, 50(9)*, 741-749.

Mezzacappa, E., & Buckner, J. C. (2010). Working memory training for children with attention problems or hyperactivity: A school-based pilot study. *School Mental Health, 2*, 202-208.

Miller, D. C. (2007). *Best practices in school neuropsychology: Guidelines for effective practice, assessment, and evidence-based intervention.* Hoboken, NJ: John Wiley & Sons, Inc.

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

Miller, G. A., Galanter, E., & Pribram, K. H. (1960). *Plans and the structure of behavior*. New York: Holt, Rinehart and Winston, Inc.

Miyake, A., Carpenter, P., & Just, M. (1994). A capacity approach to syntactic comprehension disorders: Making normal adults perform like aphasic patients. *Cognitive Neuropsychology, I f ,* 671-717.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41,* 49- 100.

Montgomery, J. W., Magimairaj, B. M., & O'Malley, M. H. (2009). Role of working memory in typically developing children's complex sentence comprehension. *Journal of Psycholinguistis, 37*, 331-354.

Montgomery, J. W., Polunenko, A., & Marinellie, S. A. (2009). Role of woring memory in understanding spoken narrative: A preliminary investigation. *Applied Psycholinguistics, 30*, 485-509.

Morey, C. C., & Cowan, N. (2004). When visual and verbal memories conflict? The importance of working- memory load and retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 703-713.

Morgan, P. L., Farkas, G., Hillemeier, M. M., & Maczuga, S. (2009). Risk factors for learning-related behavior problems at 24 months of age: Population-based estimates. *Journal of Abnormal Child Psychology, 37*, 401-413.

Muijs, D., Harris, A., Chapman, C., Stoll, L., & Russ, J. (2009). Improving schools in socioeconomically disadvantaged areas—A review of research evidence. *School Effectiveness and School Improvement, 15***,** 149-175.

Nagy, Z., Westberg, H., & Klingberg, T. (2004). Maturation of white matter is associated with the development of cognitive functions during childhood. *Journal of Cognitive Neuroscience, 16(7)*, 1227-1233.

Nevo, E., & Breznitz, Z. (2013). Assessment of working memory components at 6 years of age as predictors of reading achievements a year later. *Journal of Experimental Child Psychology, 109*, 73-90.

Norman, D. A. (1968). Toward a theory of memory and attention. *Psychological Review, 75*, 522-536.

Normand, S., & Tannock, R. (2012). Screening for working memory deficits in the classroom: The psychometric properties of the working memory rating scale in a longitudinal school-based study. *Journal of Attention Disorders, XX(X)*, 1-11.

Oberauer, K. (2009). Design for working memory. *Psychology of Learning & Motivation, 51*, 45-100.

Olive, T., & Passerault, J. (2012). The visuospatial dimension of writing. *Written Communication, 29(3)*, 326-344.

Ottem, E. J., Lian, A., & Karlsen, P. J. (2007). Reasons for the growth of traditional memory span across age. *European Journal of Cognitive Psychology,19(2)*, 233-270.

Pashler , H. E. (1984). Processing stages in overlapping tasks: Evidence for a central bottleneck. *Journal of Experimental Psychology: Human Perception and Performance, 10*, 358-377.

Pelphrey, K. & Reznick, S. (2002). Working memory in infancy. *Advances in Child Development and Behavior, 31*, 173-227.

Pelphrey, K., Reznick, S., Goldman, B., Sasson, N. Morrow, J., Donahoe, A. et al., (2004). Development of visuospatial short-term memory in second half of the 1st year. *Developmental Psychology, 40*, 836-851.

Peverly, S. T., Ramaswamy, V., Brown, C., Sumowski, J., Alidoost, M., & Garnder, J. (2007). What predicts skill in lecture note taking? *Journal of Educational Psychology, 99(1)*, 167-180.

Pickering, S. J. (2006). Working memory in dyslexia. *Working Memory and Neurodevelopmental Disorders, 7,* 5-40.

Piolat, A., Olive, T., & Kellogg, R. T. (2005). Cognitive effort during note taking. *Applied Cognitive Psychology, 19*, 291-312.

Raiker, J. S., Rapport, M. D., Kofler, M. J., & Sarver, D. E. (2012). Objectively-measured impulsivity and attention-deficit/hyperactivity disorder (ADHD): Testing competing predictions from working memory and behavioral inhibition models of ADHD. *Journal of Abnormal Child Psychology, 40*, 699-713.

Reznick, J. S., Morrow, J. D., Goldman, B. D., & Snyder, J. (2004). The onset of working memory in infants. *Infancy, 6(1)*, 145-154.

Rice, M. E. & Harris, G. T (1995). Comparing effect sizes in follow-up studies: ROC area, cohen *r,* and *d*. *Law and Human Behavior, 29(5)*, 615-620.

Rogers, M., Hwan, H., Toplak, M., Weiss, M., & Tannock, R. (2011). Inattention, working memory, and academic achievement in adolescents referred for attention deficit/hyperactivity disorder (ADHD). *Child Neuropsychology, 17(5)*, 444-458.

Salvia, J. & Ysseldyke, J. E. (2004). *Assessment in special and inclusive education.* (9th ed.) Boston: Houghton Mifflin.

Savage, R., & Lavers, N., & Pillay, V. (2007). Working memory and reading difficulties: What we know andwhat we don't know about the relationship. *Educational Psychological Reviews, 19*, 185-221.

Scherf, K. S., Sweeney, J. A., & Luna, B. (2006). Brain basis of developmental change in visuospatial working memory. *Journal of cognitive Neuroscience, 18(7),* 1045-1058.

Seigneuric, A., Ehrlich, M.-F., Oakhill, J. V., & Yuill, N. M. (2000). Working memory resources and children's reading comprehension. *Reading and Writing, 13,* 81–103.

Shallice, T., & Warrington, E. K. (1970). Independent functioning of verbal memory stores: A neuropsychological study. *Quarterly Journal of Experimental Psychology, 22*, 261-273.

Shinn, M. R. (1998). *Advanced applications of curriculum-based measurement*. New York, NY: Guildford Press.

Smith, E. E., & Jonides, J. (1997). Working memory: A view from neuroimaging. *Cognitive Psychology, 33,* 5-42.

Smith, E. E., & Jonides, J. (1999). Storage and executive processes in the frontal lobes. *Science, 283,* 1657-1661.

Sonuga-Barke, E. J., Koerting, J., Smith, E., McCann, D. C., et al. (2011). Early detection and intervention for attention-deficit/hyperactivity disorder. *Expert Review of Neurotherapeutics* 11(4):557-63.

Speece, D. L., Mills, C., & Ritchey, K. Initial evidence that letter fluency tasks are valid indicators of early reading skill. *Journal of Special Education, 36(4),* 223-233.

Spooner, D. M. & Pachana, N. A. (2006) Ecological validity in neuropsychological assessment: A case for greater consideration in research with neurologically-intact populations. *Neuropsychology Review, 21,*280-291.

St Clair-Thompson, H. L., Stevens, R., Hunt, A., & Bolder, E. (2010). Improving children's working memory and classroom performance. *Educational Psychology, 30(2),* 203-219.

St. Clair-Thompson, H. L. (2011) Executive functions and working memory behaviours in children with poor working memory. *Learning and Individual Differences, 21,* 409-414.

St. Clair-Thompson, H., L., & Gathercole, S. E. (2006). Executive functions and achievements in school: Shifting, updating, inhibition, and working memor. *The Quarterly Journal of Experimental Psychology, 59(4),* 745-759.

Stevanovski, B., & Jolicoeur, P. (2007). Visual short-term memory: Central capacity limitations in short-term consolidation. *Visual Cognition*, 15, 532-563.

Sung, J. E., McNeil, M. R., Pratt, S. R., Dickey, M. W., Hula, W. D., Szuminsky, N. J., & Doyle, P. J. (2009). Verbal working memory and its relationship to sentence-level reading and listening comprehension in persons with aphasia. *Aphasiology, 23(7-8),* 1040-1052.

Swanson, H. L. (1993). Working memory in learning disability subgroups. *Journal of Experimental Child Psychology, 56*, 87-114.

Swanson, H.L. & O'Connor, R. (2009). The role of working memory and fluency practice on the reading comprehension of students who are dysfluent readers. *Journal of Learning Disabilities,42(6),* 548-575.

Swanson, H.L. (2011). Working memory, attention, and mathematical problem solving: A longitudinal study of elementary school children. *Journal of Educational Psychology, 103(4)*, 821-837.

Sweller, J., van Merrienboer, J., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*, 251-296.

Taub, G. E. & McGrew, K. S. (2014). The woodcock-johnson tests of cognitive abilities III's cognitive performance model: Empirical support for intermediate factors within CHC theory. *Journal of Psychoeducational Assessment, 32(3)*, 187-201.

Thurman, S. K., & Kiepert, M. H. (2008). Issues and concerns in conducting applied cognitive research. In S. K. Thurman, & C. A. Fiorello (Eds.) *Applied cognitive research in K-3 classrooms* (pp. 265-288) New York, NY: Routledge/Taylor & Francis Group.

Treisman, A. M. (1960). Contextual cues in selective listening. *Quarterly Journal of Experimental Psychology, 12*, 242-248.

United States Census Bureau (2013). State and County QuickFacts. Data derived from Population Estimates, American Community Survey, Census of Population and Housing, State and County Housing Unit Estimates, County Business Patterns, Nonemployer Statistics, Economic Census, Survey of Business Owners, Building Permits. Last Revised: Tuesday, 08-Jul-2014 06:37:33 EDT

Vallar, G., & Baddeley, A. (1984). Fractionation of working memory: Neuropsychological evidence for a phonological short-term store. *Journal of Verbal Learning and Verbal Behavior, 23*, 151-161.

Warrington, E. K., & Shallice, T. (1969). The selective impairment of auditory verbal short-term memory. *Brain*, *92*, 885-896.

Wass, S. V., Scerif, G., & Johnson, M. H. (2012) Training attentional control and working memory- Is younger better? *Developmental Review (32)*, 360-387.

Weschler, D. (2003). *Wechsler Intelligence Scale for Children- Fourth Edition*. San Antonio, Texas: Pearson.

Wright, H., Downey, R. A., Gravier, M., Love, T., & Shapiro, L. P. (2007). Processing distinct linguistic informatino types in working memory in aphasia. *Aphasiology, 21(6-8)*, 802-813.

Wu, K. K., Chan, S. K., Leung, P. W. L., Liu, W., Leung, F. L. T., & Ng, R. (2011). Components and developmental differences of executive functioning for school-ages children. *Developmental Neuropsychology, 36(3)*, 319-337.

Zheng, X., Swansaon, H.L., & Marcoulides, G.A. (2011). Working memory components as predictors of children's mathematical word problem-solving. *Journal of Experimental Child Psychology, 110,* 481-498.

APPENDIX A

PARENT CONSENT FORM

Title of research study:
Construction of a Working Memory Screener for Elementary School Children

Investigator and Department:
Principal Investigator:
Catherine Fiorello, Ph.D.
Associate Professor and Coordinator of School Psychology Program
Department of Psychological, Organizational, & Leadership Studies in Education
College of Education, Temple University

Why you are being invited to take part in a research study
We invite you and your child to take part in a research study because you are a student enrolled in a regular /inclusion elementary school classroom within the greater Philadelphia area. .

Who can I talk to?
If you have questions, concerns, or complaints, or think the research has hurt you or your child, contact the research team at
Catherine Fiorello, Ph.D, NCSP
Ritter Annex 269
1301 Cecil B. Moore Avenue
Philadelphia, PA 19122-6091
catherine.fiorello@temple.edu
215-204-6254

Gina Forchelli, Ed.M. NCSP
Gina.Forchelli@temple.edu

This research has been reviewed and approved by an Institutional Review Board. You may talk to them at (215) 707-3390 or e-mail them at: irb@temple.edu for any of the following:
Your questions, concerns, or complaints are not being answered by the research team.
You cannot reach the research team.
You want to talk to someone besides the research team.
You have questions about your rights as a research subject.
You want to get information or provide input about this research.
Why are we doing this research?

Working memory, or the ability to hold and transform information in one's head, has been linked to academic performance in school-aged children. Preliminary research suggests that, with remediation, working memory skills can improve and can sometimes transfer to improvement in academic performance. However, identification methods of working memory difficulty are limited. It is my hope to develop a measure that is easy to administer and identify children with working memory difficulties in order to facilitate remediation.

How long will the research last?

We expect that you and your child will be in this research study for one to three months.

How many people will be studied?

We expect about 30 people here will be in this research study out of 60 people in the entire study nationally.

What happens if I say yes, I want to be in this research?

You, the parent, will complete a brief demographic questionnaire and a form rating your child's behavior over the past 6 months. These can be returned to your child's teacher or mailed back to the principal investigators address directly. Your child's teacher will also be asked to complete a similar form rating on your child's behavior over the past 6 months.

Your child will be asked to meet briefly with the principal investigator or graduate level student during class time to complete a short cognitive task; this will approximately take 10 minutes. Your child will also be asked to participate in a 10-minute group-administered task during class-time. After your child completes these tasks, you, the parent, and your child's teacher will get a brief report on how your child performed on the individually-administered cognitive tasks to help you understand more on how your child learns information.

What are my responsibilities if I take part in this research?

If you take part in this research, you and your child will be responsible for actively participating in the tasks required and providing truthful information on rating and demographic forms.

What happens if I say no, I do not want to be in this research?

You and your child may decide not to take part in the research and it will not be held against you. It will in no way affect you and your child's relationship with your school or your teacher.

What happens if I say yes, but I change my mind later?

If you and your child agree to take part in the research now and if you and your child stop at any time, it will not be held against you. Again, it will in no way affect your relationship with your child's school or teacher. If you and your child stop participating in the research, already collected data may not be removed from the study database.

Is there any way being in this study could be bad for me?

There is minimal risk associated with participation with this study. You and your child may become somewhat fatigued after the completion of each task; however, no other risk or discomfort is expected. Your child will miss a small amount of instruction (approximately 10 minutes) to complete the individually administered tasks.

Will being in this study help me any way?

The possible benefits include you and your child's' increased awareness of your child's memory and learning skills and you will also be contributing to the development of a measure that may aid students in the future.

What happens to the information we collect?

Efforts will be made to limit you and your child's personal information to people who have a need to review this information. All participant data will be entered into a database with codes in place of names by the PI. The original scored forms will be stored in a locked filing cabinet in the PI's department office, and all analyses will be completed with the de-identified data. The PI will complete a results form to return to the district evaluation team and one for the school to send to the parent. We cannot promise complete secrecy, as there is always a potential risk of loss of confidentiality. Organizations that may inspect and copy your information include the IRB, Temple University, Temple University Health System, Inc. and its affiliates, and other representatives of these organizations, and the Office of Human Research Protections. We may publish the results of this research. However, we will keep you/your child's name and other identifying information confidential.

Can I be removed from the research without my permission?

The person in charge of the research study can remove you and your child from the research study without your approval.

We will tell you about any new information that may affect your health, welfare, or choice to stay in the research.

What else do I need to know?

If you/your child sustain an injury as a result of your participation in this research study, the physician's fees and medical expenses that result will be billed to your insurance company or you in the usual manner. Other financial compensation (such as lost wages or pain and suffering) for such injuries is not routinely available. By signing this consent form, you are not waiving any of the legal rights that you otherwise would have as a participant in a research study. If you have questions about the study or a research-related injury, please contact calling Dr. Fiorello at 215-204-6254 or e-mailing catherine.fiorello@temple.edu.

Stipend/Reimbursement

There is no stipend associated with participation.

Participating in Future Research Studies

We may want to contact you in the future to see if you would be interested in participating in another research study and/or to obtain additional information related to your participation in this study. Please indicate by initialing on the line in the next paragraph below if you are willing to be contacted. Please know that you can amend your answer below at any time without prejudice to you or your relationship with the study, Temple University, or the team.

	Initial your choices
Yes, I agree to be contacted about future research studies.	______
OR	
No, I do not want to be contacted about future research studies.	______
Yes, I agree to be contacted to obtain additional information related to my participation in this study.	______
OR	
No, I do not want to be contacted to obtain additional information related to my participation in this study.	______

Your signature below indicates that:
Someone has explained this research study to you.
You freely volunteer to be in this research study.
You can choose not to take part in this research study and it will not affect your care.
You can agree to take part in this study now and later change your mind. Your decision to leave the study will not affect your care.
You have been offered the opportunity to ask questions and all your questions have been answered.

Your signature documents your permission for the named child to take part in this research.

DO NOT SIGN THIS FORM AFTER THIS DATE →

Printed name of child

Signature of parent or guardian | Date

☐ Parent

Printed name of parent or guardian

☐ Guardian (See note below)

Note on permission by guardians: An individual may provide permission for a child only if that individual can provide a written document indicating that he or she is legally authorized to consent to the child's general medical care. Attach the documentation to the signed document.

Signature of person obtaining consent and assent | Date

Printed name of person obtaining consent and assent | Date

APPENDIX B

PARTICIPANT ASSENT FORM

My name is [researcher name]. I am trying to learn about how kids learn and remember information. It will help teachers and adults understand how to teach kids in better ways to help you do well in school. If you would like, you can be in my study.

If you decide you want to be in my study, you will be asked to complete two tasks with me one-on-one. In one, I will ask you to remember some numbers and letters for me and repeat them back to me. In another, I will ask you to play with some blocks. After we complete these, I will also ask you to complete a task with your whole class. You will follow a list of directions that are asked of you.

If you participate, you will get to understand more about how you understand and learn. You will also will help adults teach kids better in the future. You may get a little tired after the completion of each task; however, this should be very little amount.

Other people will not know if you are in my study. I will put things I learn about you together with things I learn about other students, so no one can tell what things came from you. When I tell other people about my research, I will not use your name, so no one can tell who I am talking about.

Your parents or guardian have to say it's OK for you to be in the study. After they decide, you get to choose if you want to do it too. If you don't want to be in the study, no one will be mad at you. If you want to be in the study now and change your mind later, that's OK. You can stop at any time.

My telephone number is 845-641-6962. You can call me if you have questions about the study or if you decide you don't want to be in the study any more.

I will give you a copy of this form in case you want to ask questions later.

Agreement

I [Name of Researcher] am confirming that [Name of Student] verbally assented to proceed with the reviewed study.

________________________________ _________________

Signature of Researcher Date

APPENDIX C

TEACHER CONSENT FORM

Title of research study:
Construction of a Working Memory Screener for Elementary School Children

Investigator and Department:
Principal Investigator:
Catherine Fiorello, Ph.D.
Associate Professor and Coordinator of School Psychology Program
Department of Psychological, Organizational, & Leadership Studies in Education
College of Education, Temple University

Why you are being invited to take part in a research study
We invite you to take part in a research study because you are a teacher of a regular /inclusion elementary school classroom within the greater Philadelphia area.

Who can I talk to?
If you have questions, concerns, or complaints, or think the research has hurt you, contact the research team at
Catherine Fiorello, Ph.D, NCSP
Ritter Annex 269
1301 Cecil B. Moore Avenue
Philadelphia, PA 19122-6091
catherine.fiorello@temple.edu
215-204-6254

Gina Forchelli, Ed.M. NCSP
Gina.Forchelli@temple.edu

This research has been reviewed and approved by an Institutional Review Board. You may talk to them at (215) 707-3390 or e-mail them at: irb@temple.edu for any of the following:

- Your questions, concerns, or complaints are not being answered by the research team.
- You cannot reach the research team.
- You want to talk to someone besides the research team.
- You have questions about your rights as a research subject.
- You want to get information or provide input about this research.

Why are we doing this research?
Working memory, or the ability to hold and transform information in one's head, has been linked to academic performance in school-aged children. Preliminary research suggests that, with remediation, working memory skills can improve and can sometimes transfer to improvement in academic performance. However, identification methods of working memory difficulty are limited. It is my hope to develop a measure that is easy to administer and identify children with working memory difficulties in order to facilitate remediation.

How long will the research last?
We expect that you will be in this research study for one to three months.
How many people will be studied?

We expect about 30 students here will be in this research study out of 60 people in the entire study nationally.

What happens if I say yes, I want to be in this research?
You will be asked to complete three tasks during the study. First, you will complete rating forms for the participating students' behavior over the past 6 months. Second, you will be asked to allow the principal investigator or graduate level to meet with students during class time to complete a short memory task. Last, you will be asked to administer or allow the principal investigator to administer a 10-minute group-administered task during class-time. More specifically, you will pass out answer sheets for a following directions task, play a recording that has all directions and test items, and collect the papers at completion. At the conclusion of these tasks, you will receive a brief report indicating how the participating students performed on standardized measures administered.

What are my responsibilities if I take part in this research?
If you take part in this research, you will be responsible for providing truthful information on rating forms and collecting papers during the classroom-wide following directions task.

What happens if I say no, I do not want to be in this research?
You may decide not to take part in the research and it will not be held against you. It will in no way affect your relationship with your school or your students.

What happens if I say yes, but I change my mind later?
You agree to take part in the research now and if you stop at any time, it will not be held against you. Again, it will in no way affect your relationship with your school or your students. If you stop being in the research, already collected data may not be removed from the study database.

Is there any way being in this study could be bad for me?
There is no anticipated risk associated with your participation with this study.
Will being in this study help me any way?
The benefits will be an increased awareness of your students' memory and learning skills and you will also be contributing to the development of a measure that may aid students in the future.

What happens to the information we collect?
Efforts will be made to limit your personal information to people who have a need to review this information. All participant data will be entered into a database with codes in place of names by the PI. The original scored forms will be stored in a locked filing cabinet in the PI's department office, and all analyses will be completed with the de-identified data. The PI will complete a results form to return to the district evaluation team and one for the school to send to the parent. We cannot promise complete secrecy, as there is always a potential risk of loss of confidentiality. Organizations that may inspect and copy your information include the IRB, Temple University, Temple University Health System, Inc. and its affiliates, and other representatives of these organizations, and the Office of Human Research Protections. We may publish the results of this research. However, we will keep your name and other identifying information confidential.

Can I be removed from the research without my permission?
The person in charge of the research study can remove you from the research study without your approval. The sponsor can also end the research study early.
We will tell you about any new information that may affect your health, welfare, or choice to stay in the research.

What else do I need to know?

If you sustain an injury as a result of your participation in this research study, the physician's fees and medical expenses that result will be billed to your insurance company or you in the usual manner. Other financial compensation (such as lost wages or pain and suffering) for such injuries is not routinely available. By signing this consent form, you are not waiving any of the legal rights that you otherwise would have as a participant in a research study. If you have questions about the study or a research-related injury, please contact calling Dr. Fiorello at 215-204-6254 or e-mailing catherine.fiorello@temple.edu.

Stipend/Reimbursement

There is no stipend associated with participation.

Participating in Future Research Studies

We may want to contact you in the future to see if you would be interested in participating in another research study and/or to obtain additional information related to your participation in this study. Please indicate by initialing on the line in the next paragraph below if you are willing to be contacted. Please know that you can amend your answer below at any time without prejudice to you or your relationship with the study, Temple University, or the team.

Initial your choices

Yes, I agree to be contacted about future research studies. ____________

OR

No, I do not want to be contacted about future research studies. ____________

Yes, I agree to be contacted to obtain additional information related to my participation in this study. ____________

OR

No, I do not want to be contacted to obtain additional information related to my participation in this study. ____________

Your signature below indicates that:

- Someone has explained this research study to you.
- You freely volunteer to be in this research study.
- You can choose not to take part in this research study and it will not affect your care.
- You can agree to take part in this study now and later change your mind. Your decision to leave the study will not affect your care.
- You have been offered the opportunity to ask questions and all your questions have been answered.

Your signature documents your permission to take part in this research.

DO NOT SIGN THIS FORM AFTER THIS DATE →

____________________ ____________

Signature of subject | Date

____

Printed name of subject

____________________ ____________

Signature of person obtaining consent	Date
Printed name of person obtaining consent	

APPENDIX D

PARENT DEMOGRAPHIC QUESTIONNAIRE

1. Name:____________________
2. Birthday: ____________________
3. Gender: Male or Female
4. Race: European-American, African-American, Asian-American, Other
    a. Hispanic descent?
5. School:
6. Grade:
7. Languages spoken in the home:
    a. Primary language:
    b. Secondary language:
8. Currently receiving special education services:
    a. What diagnosis?
    b. Other diagnoses?

APPENDIX E

FORCHELLI FOLLOWING DIRECTIONS TASK©2014

Directions/Answer Key:

Administrator instructions: Before beginning the task, check to make sure the audio recording is loud enough so that the students can comfortably hear the items. Please provide each student with a response sheet face down on their desk and provide each student with a black pen. After every student has their materials and all students are seated, read the following prompt:

"We are going to play a game! I am going to have you listen to a list of directions from this recording. Please listen to each direction and after I say 'Go,' follow what the speaker asks you to do. I cannot repeat any of the directions on the recording so please listen carefully. Ready?"

Start playing the audio recording and follow along with each prompt. At the completion of number 15, stop the audio recording and collect the participants' responses.

Scoring. Participants receive 1-point for each behavior correctly followed on their response sheet. Possible points are indicated next to each item below. For all items, number and letter reversals still receive full credit. However, order of letters and numbers matter.

Directions

(Play by audio recording 1 word/second and 5 second pause between directions)

Turn your paper over. Go. (1 behavior)

1. Draw a circle next to number 1. Go.

(1 behavior + 2 qualifiers)

Scoring: +1 point for circle next to number 1

2. Draw a square next to number 2 only if you are standing up. Go.

(1 behavior + 3 qualifiers)

Scoring: +1 point for blank on number 2. No credit received if erased.

3. Print the first letter of your name next to number 3. Go.

(1 behavior + 3 qualifiers)

Scoring: +1 point for first letter of name. No credit if they write their full name.

4. Draw a square and draw a circle on number 4. Go.

(2 behaviors + 1 qualifier)

Scoring: +1 point square on number 4, +1 point for circle on number 4

5. Next to number 5, circle the third dot. Go.

(2 behaviors + 2 qualifiers)

Scoring: +2 points for ONLY circling third dot

6. Write the letter "O" only after you write the letter "G" next to number 6.

(2 behaviors + 3 qualifiers)

Scoring: +1 point for G on number 6, +1 point for O on number 6, +1 point for correct order

7. Write the number 8 next to the circle on number 1. Go.

(2 behaviors + 4 qualifiers)

Scoring: +1 point for drawing an "8" on the page, + 2 points if next to the circle on number 1

8. Color in the first star next to number 7 and then color in third star. Go.

(2 behaviors + 4 qualifiers).

Scoring: +1 point for colored first star on number 7, +1 point for colored third star, +1 if both correctly colored. Still receive points if also colored in an additional star.

9. Write three letters of the alphabet next to number 8. Go.

(3 behaviors + 1 qualifier)

Scoring: +1 point for each letter (max 3) next to number 8

10. On number 9, draw an "X" only after you draw a square. Go.

(3 behaviors + 2 qualifiers)

Scoring: +1 point for "X" next to number 9, +1 point for square next to number 9, +1 point for correct sequence

11. On number 10, put an "X" in the second box then put an "X" in the first box and the last box. Go.

(3 behaviors + 3 qualifiers)

Scoring: +1 point for "X"ed first box on number 10, +1 point for X"ed second box on number 10, +1 point for X"ed last (fifth) box on number 10, +1 point for all correct

12. Add 1 and 1 and 2 and write down the answer next to number 11. Go.

(3 behaviors + 3 qualifiers)

Scoring: +2 points for writing "4" next to number 11, +1 point if they DID NOT write down the equation

13. On number 12, circle the word cat then circle the word mat and circle a word that rhymes with bed. Go.

(4 behaviors + 2 qualifiers)

Scoring: +1 point cat circled, +1 point for mat circled, +1 point for red circled, +1 point for all correct.

14. On number 13, listen to the following sentence and write the second word and write the last word and write the word "rat". *The cow is fat.* Go.

(4 behaviors + 3 qualifiers)

Scoring: +1 point for cow, +1 point for fat, +1 point for rat, +1 for all correct)

15. On number 14, draw the number 3, then draw a second number 3 only after drawing a plus sign. Finally, then draw an equals sign. Go.

(4 behaviors + 4 qualifiers)

Scoring: +1 point for drawing 3, +1 point for drawing second 3, +1 point drawing plus sign, +1 point drawing equal sign, +1 point for correct sequence (3+3=)

Forchelli Following Directions Task©2014

Response Sheet

Name: ____________

Date: ____________

Total Score: ____/42____

1. ____________

2. ____________

3. ____________

4. ____________

5. ● ● ● ● ● ●

6. ____________

7. ☆ ☆ ☆ ☆ ☆

8. ____________

9. ____________

10. □ □ □ □ □

11. ____________

12. The cat and the mat are red.

13. ____________

14. ____________

APPENDIX F

STUDENT REPORT FORM

WISC-IV Working Memory Subtest Scores

Cognitive Battery/Subtest		**Scaled/Standard Score**	**Classification**
Digit Span	**Digit Span Forward**		
	Digit Span Backwards		
Spatial Span	**Spatial Span Forward**		
	Spatial Span Backwards		

Behavioral Inventory of Executive Functions (BRIEF)

	Parent Rating Form T-Score/ Classification	**Teacher Rating Form T-Score/ Classification**
Emotional Control		
Initiate		
Organization of Materials		

Plan/Organize		
Self-Monitor		
Shift		
Working Memory		

Date of Report: _____________

Compiled by: _____________________________________

Gina Forchelli, Ed.M., NCSP

Temple University

APPENDIX G

BRIEF PARENT/TEACHER RATING FORM SUMMARIES

Behavioral Regulation Index of Executive Functions (BRIEF)- Teacher and Parent forms (Gioia, Isquith, Guy, & Kenworthy, 2009). The BRIEF is a 86-item behavior checklist that measures various abilities considered within the executive functions construct. Raters assess the child's behavior on a 3-point Likert-like scale (i.e., Never, Sometimes, or Often). Raw totals are obtained and converted into T-scores for Emotional Control, Initiate, Inhibit, Organization of Materials, Plan/Organize, Self-Monitor, Shift, and Working Memory. Factor scores are also provided for the Metacognition and Behavioral Regulation indexes, as well as a Global Executive Composite. The manual indicates that higher T-scores (i.e., above a score of 60) indicate more difficulty within an area. The Working Memory Index will be the only scale analyzed in this study. Sample question from each section include:

Inhibit: "Needs to be told 'no' or 'stop that'"

Shift: "Cannot get a disappointment, scolding, or insult off his/her mind"

Emotional Control: "Overreacts to small problems"

Initiate:"Is not a self-starter"

Working Memory: "When given three things to do, remembers only the first or last"

Plan/Organize:"Does not bring home homework, assignment sheets, materials, etc."

Organization of Materials:"Loses lunch box, lunch money, permission slips, homework, etc."

Monitor: "Does not check work for mistakes"

APPENDIX H

WISC-IV FORWARD AND BACKWARD DIGIT SPAN AND SPATIAL SPAN SUMMARIES

Digit Span Forward (DSF) and Backwards (DSB) (Weschler, 2003). Digit Span is an individually administered objective measure of verbal WM. In the forward condition of this measure, participants hear a sequence of digits at a rate of one digit per second and are asked to immediately repeat the sequence in the exact order it was presented. The length of the sequence begins with two digits and becomes increasingly more difficult (up to nine digits in length) until the participant receives two errors in a set. In the backwards condition, participants hear a sequence of digits at a rate of one digit per second and are asked to repeat the sequence in the reverse order that is was presented. Similar to DSF, DSB begins with two digits in length and increases to nine digits in length.

Forwards and Backwards Spatial Span (Weschler, 2003). Spatial Span is an individually administered subtest from the WISC-IV integrated that is thought to tap an individual's visual-spatial WM. The measure utilizes a board with randomly located blue blocks attached to it. In the forward condition, the individual touches a increasingly difficult sequence of blocks in the same order that was modeled by the examiner. In the backward condition, the individual is required to reverse the order of the block sequence modeled by the examiner. For each condition, the length of the sequence begins with two digits and becomes increasingly more difficult (up to eight digits in length) until the participant receives two errors

APPENDIX I

TEACHER RATING OF ECOLOGICAL RELEVANCE

After reviewing the instructions and items of the Following Direction Task, please answer the following questions:

1. Are the instructions simplified enough for elementary school (K-3 grade) students to understand? If no, please provide a comment below.

2. Do the items reflect classroom activities that your students engage in? If not, please specify which items.

3. Are any of the items too difficult for your students? If yes, please describe below.

APPENDIX J

EXPERT RATING OF WORKING MEMORY MEASUREMENT

After reviewing the instructions and items of the Following Direction Task, please answer the following questions:

1. Are the items within the questionnaire representative of working memory abilities? If no, please describe below.

2. Do you have suggestions on items or tasks that should be added to the questionnaire to enhance the detection of working memory skill?