

# UC Riverside

## UC Riverside Electronic Theses and Dissertations

### Title

Age Related Differences in the Dynamic Assessment of Working Memory When Predicting Reading and Math Outcomes

### Permalink

<https://escholarship.org/uc/item/6kp7v3d8>

### Author

Saunders, Jared J.

### Publication Date

2018

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA  
RIVERSIDE

Age Related Differences in the Dynamic Assessment of Working Memory When  
Predicting Reading and Math Outcomes

A Dissertation submitted in partial satisfaction  
of the requirements for the degree of

Doctor of Philosophy

in

Educational Psychology

by

Jared James Saunders

March 2018

Dissertation Committee:

Dr. H. Lee Swanson, Chairperson

Dr. Rollanda O'Connor

Dr. Gregory Palardy

Copyright by  
Jared James Saunders  
2018

The Dissertation of Jared James Saunders is approved:

---

---

---

Committee Chairperson

University of California, Riverside

## ABSTRACT OF THE DISSERTATION

Age Related Differences in the Dynamic Assessment of Working Memory When  
Predicting Reading and Math Outcomes

by

Jared James Saunders

Doctor of Philosophy, Graduate Program in Educational Psychology  
University of California, Riverside, March 2018  
Dr. H. Lee Swanson, Chairperson

Traditional approaches to intelligence and cognitive assessment have been criticized because they assess an individual's prior knowledge rather than their aptitude to learn. Dynamic assessment has been proposed to address this limitation by integrating forms of learning into the assessment process. Dynamic assessment has been applied to many different arenas including intelligence, speech and language, and areas of achievement, and has generally been found to predict additional variance in the criterion measure beyond that which is predicted by the static measures. However, the variables that potentially moderate this additional variance have not been clearly explored. The purpose of this study is to evaluate at least three moderating variables (age, modality, type of dynamic measure) that may interact with the contribution of dynamic assessment of working memory performance in the predictions of reading and math performance. Three research questions are proposed which ask (1) does dynamic assessment of working memory contribute unique variance in predictions of reading and math performance above that of the static assessment, (2) does age moderate the effectiveness

of dynamic assessment of a domain general construct of working memory in predictions of achievement measures, and (3) does performance on visual and verbal measures of working memory interact with age when predicting reading and math performance.

Results of the study find support that dynamic assessment contributes unique variance in predictions of reading and math achievement. However, these results were qualified since age x dynamic interactions emerged when analyzing working memory as both a domain general and domain specific construct. The majority of these interactions were limited to predictions of math achievement. Implications of these findings and limitations of the study are discussed.

## Table of Contents

Chapter 1:	Rationale and Significance	1
Chapter 2:	Literature Review	17
Chapter 3:	Methods	34
Chapter 4:	Results	44
Chapter 5:	Discussion	61
References:		75
Appendix:		85

## **List of Figures**

Figure 1: Math Performance as a Function of WM Gain Scores	57
Figure 2: Math Performance as a Function of WM Maintenance Scores	59



## **List of Tables**

Table 1: Characteristics of the Sample	85
Table 2: Correlations Between Independent and Dependent Variables	86
Table 3: Dynamic Assessment of a Domain General Construct of WM on Reading and Math	87
Table 4: Dynamic Assessment of WM by Modality on Reading and Math	88
Table 5: Descriptive Statistics for the Dependent Variable by Age Group	89
Table 6: ANCOVA Comparisons of Age x Gain Scores When Predicting Math Achievement	90
Table 7: ANCOVA Comparisons of Age x Maintenance Scores When Predicting Math Achievement	91

## **Chapter 1 - Rationale and Significance**

Several problems have been noted with the use of static, traditional assessment to measure potential in children. Traditional assessments only serve as a measure of what the child has already learned rather than a measurement of his or her true potential to learn. Haywood and Tzuriel (2002) pointed out that traditional assessment is able to effectively predict 70-75% of the variance in academic achievement, leaving 25-30% of performance unaccounted for. With such a large remaining gap in academic outcome, several researchers recommended that dynamic assessment practices be added as a supplement to, rather than a replacement of, traditional, static assessment practices to improve prediction ability (Murphy & Maree, 2006).

In addition, Carlson and Wiedl (1992) indicated environmental validity is enhanced with the use of dynamic assessment. This is because intervention has a close relationship with dynamic assessment as identifying appropriate accommodations and strategies is often one of its primary purposes (Embretson, 1987; Grigorenko & Sternberg, 1998). Dynamic assessment is suited to identify effective interventions for an individual because it seeks to identify whether a student's performance on a test is able to change when support is provided, and to evaluate how that change occurs and what it looks like (Grigorenko, 2009; Grigorenko & Sternberg, 1998).

Although several definitions and variations of dynamic assessment exist, it is generally viewed as a process in which active thinking, perception, problem solving, and learning is measured within the context of an instructional paradigm that seeks to modify cognition (Haywood & Tzuriel, 2002). In addition, other commonalities between

methods of dynamic assessment include the role of the examiner within the assessment process as being active and consisting of an interaction that occurs between the examiner and examinee as opposed to a one-way questioner-responder dyad (Haywood & Wingenfeld, 1992). In dynamic assessment, there is an intentional effort to alter the examinee's performance on the test, with a more general goal of assessing a degree of potential rather than accumulated knowledge as it exists at the time of assessment.

Haywood and Tzuriel (2002) identified several assumptions that underlie dynamic assessment. First, what an individual has learned in the past is not the best indicator of what an individual can and will learn in the future. A second assumption is that all individuals perform below their optimal level; therefore, everyone can improve with help. Third, it is assumed that the best way to test learning is to obtain samples of learning in progress, rather than sampling what has already been learned. A fourth assumption is that people perform below their potential due to a number of obstacles (e.g., motivation); however, if these obstacles are removed, improvements in performance may be observed.

### **Origins in Vygotsky**

Although many psychologists, such as Binet, Thorndike, and Rey, have independently made reference to the use of testing procedures that resemble dynamic assessment, Lev Vygotsky (1978) is generally described as its founder (Campione & Brown, 1987; Grigorenko & Sternberg, 1998; Haywood & Tzuriel 2002; Murphy & Maree, 2006). In his framework, Vygotsky identified two zones of performance: the zone of current development, and the zone of proximal development (ZPD). The zone of current development is defined as the level of performance an individual can reach

independently, while the ZPD refers to the difference between what an individual can do independently, and what he or she can do when provided assistance by a more experienced learner or teacher (Guthke & Wingenfeld, 1992; Haywood & Tzuriel, 2002). The ZPD is where development and maturation take place. Because this development occurs as a result of interactions between persons, assessment of the child within the context of interaction is necessary in order to measure the ZPD (Murphy & Maree, 2006). It has been predicted that the larger a child's ZPD is, the more success that individual will likely experience in school (Campione & Brown, 1987; Grigorenko & Sternberg, 1998; Haywood & Tzuriel, 2002). Although many researchers in the past four decades have developed models of dynamic assessment, Vygotsky never actually empirically validated this theory, nor did he provide guidelines as to how the ZPD should be measured (Guthke & Wingenfeld, 1992; Tudge & Scrimsher, 2003). Vygotsky did identify these guidelines as the next step in his professional career, but unfortunately he died before they were realized (Tudge & Scrimsher, 2003).

Campione and Brown (1987) pointed out that assessment of the ZPD should not be viewed as a stable trait, as differences between what a child can do independently and what he or she can do with a little help is not consistent over time. Instead, they recommend that assessment be ongoing in order to capture changes of the zone, as each measurement is only meaningful for a brief period of time. In order to optimize transfer of learning to other environments, Campione and Brown (1987) recommended attention be paid to the metacognitive factors within the instructional setting by ensuring students

are aware of the skills they are taught and the importance of monitoring and adjusting those skills.

### **Strengths of Dynamic Assessment**

When taken together, it appears there is some evidence that dynamic assessment can contribute to an assessment battery in meaningful ways. Swanson and Lussier (2001) conducted a meta-analysis to address some of the validity issues raised by Grigorenko and Sternberg's (1998) comprehensive review of the dynamic assessment literature. In their analysis, Swanson and Lussier (2001) tested to see if changes in testing performance extended beyond gains that occur secondary to retesting. Since Lipsey and Wilson (1993, as cited in Swanson & Lussier, 2001) reported a mean effect size of .76 across 45 different studies, Swanson and Lussier (2001) controlled for the effects of retest by subtracting this value from their findings. In addition, they wanted to know if the literature on dynamic assessment showed variations in outcomes based on treatment conditions. In total, they were able to draw the needed data to compute effect sizes from 30 articles. Based on their analysis, Swanson and Lussier found main effects for age of participants, sample size, categorization of participant disability/need, type of instruction given during testing, and type of design after correcting the weighted effect sizes as a function of categorical variables contributed to dynamic assessment outcomes.

Swanson and Lussier (2001) indicated that the use of dynamic assessment “substantially improved testing performance over static testing conditions” (p. 341), and that these changes occur above and beyond the effects that one would expect to see from retesting alone. Strategy training, provision of general feedback, and modelling of

responses produced the greatest effect sizes, followed by scaffolding procedures. They did, however, agree that the studies included in the analysis were susceptible to poor psychometric properties, were not validated alongside academic assessments, and generally focused on the psychometrically controversial change score rather than process, although many of these concerns are addressed in part by the Swanson-Cognitive Processing Test (S-CPT, 1996), which was later revised as the abbreviated Test of Working Memory (aTWM, Swanson, 2013), as well as related literature discussed in this review.

Caffrey, Fuchs, and Fuchs (2008) also evaluated the validity of applications of dynamic assessment. They conducted a systematic review of the literature to evaluate the predictive validity of dynamic assessment measures. Their review consisted of 24 studies that included individuals with high incidence disabilities, students who were at-risk for learning problems, English language learners, and typically achieving students. They found that dynamic assessment was able to predict future achievement best when the assessment provided non-contingent feedback instead of feedback reliant on student response, and when the participants were identified as having a high incidence disability rather than those identified as academically at-risk or English language learners. In addition, criterion-referenced assessments were better suited for detecting change than norm-referenced assessments. These authors concluded that dynamic assessment procedures were able to predict academic performance above and beyond that of static, traditional tests.

## **Current Trends/Domain-Specific Applications of Dynamic Assessment**

Publications in the field of dynamic assessment have tended to move away from making predictions of performance in domain general skill or aptitude areas such as intelligence, in favor of domain specific constructs such as achievement (e.g., reading, math), specific cognitive processes (e.g., planning, executive processing), and its relationship with response to intervention (Grigorenko, 2009). As an example, dynamic assessment has been applied to the field of executive functions. Studies in this field have used dynamic assessment to predict fluid intelligence scores through training in the area of planning (Cormier, Carlson, & Das, 1990), identifying strategies that improve planning (Kar, Dash Utkal, Das, & Carlson, 1993), differentiating treatment outcomes for individuals with traumatic brain injuries (Uprichard, Kupshik, Pine, & Fletcher, 2009), identifying sources of poor performance in tests of executive functioning (Kirkwood, Weiler, Bernstein, Forbes, & Waber 2001; Waber, Isquith, Kahn, Romero, Sallan, & Tarbell, 1994), and evaluating the modifiability of metacognitive skills (Saldana, 2004a; 2004b).

In the field of speech and language, Boers, Janssen, Minnaert, and Ruijssenaars (2013) reviewed the literature and found several valid uses of dynamic assessment. These included identifying environmental and contextual cues that supported communication (Nigram, 2001), determining the amount of assistance required for a child to demonstrate a target response (McLaughlin & Cascella, 2008; Snell, 2002), producing one word utterances (Gutierrez-Clellen & Pena, 2001; Lidz & Pena, 1996; Pena, Iglesias, Lidz, 2001), communicating important features of a story (Gutierrez-Clellen, Pena, & Quinn,

1995), training phonology/articulation (Glaspey & MacLeod, 2010; Glaspey & Stoel-Gammon, 2007; Spector, 1992), and requesting information (Donaldson & Olswang, 2007).

Dynamic assessment has also made significant contributions within K-12 education. In its application to phonological awareness, multiple researchers have found that dynamic assessment can more accurately predict future performance on phonological measures and other areas associated with reading than static, traditional assessment (Bridges & Catts, 2011; Coventry, Byrne, Olson, Corley, & Samuelsson, 2011; Gilliam, Fargo, Foley, & Olszewski, 2011; O’Conner & Jenkins, 1999; Spector, 1992). Dynamic assessment has been applied to better predict first grade reading ability (Peterson, Allen, & Spencer, 2014), and even differentiate between students with and without dyslexia (Aravena, Tijms, Snellings, & van der Molen, 2015, as cited in Aravena, Tijms, Snellings, & van der Molen, 2016). It has also been used to predict future responsiveness to intervention (Aravena, et al., 2016; Cho, Compton, Fuchs, Fuchs, & Bouton, 2014).

In the area of reading comprehension, dynamic assessment has been used successfully to predict a student’s ability to learn and use strategies in reading comprehension (Barabadi & Kamrood, 2011; Kozulin and Garb, 2002; Pishghadam, Barabadi, & Kamrood, 2011).

In mathematics, dynamic assessment has been used to predict novel word problem solving in third grade children (Fuchs, Compton, Fuchs, Hollenbeck, Craddock, & Hamlett, 2008) and has been demonstrated to be a useful screening tool to apply as part of a response to mathematically-based intervention model (Fuchs, Compton, Fuchs,



Hollenbeck, Hamlett, & Seethaler, 2011). Dynamic assessment has also been validated as a tool to help teach strategic math intervention, which targets word problem solving (Kong & Orozco, 2016; Orosco, Swanson, O’Conner, & Lussier, 2013).

Another application of dynamic assessment that has received some attention in the literature is the modification of working memory (e.g., Swanson, 1992). Because the application of dynamic assessment on working memory measures is the focus of this proposal, a brief context of this literature is provided below.

### **Dynamic Assessment and Working Memory**

Working memory has been defined as the ability to hold on to information while manipulating or processing similar information (Baddeley 2000; Engle, Tuholski, Laughlin, & Conway, 1999). To accomplish this aim, the working memory system draws upon diverse cognitive resources from different systems in the brain. Baddeley and Hitch (1974) developed one of the most widely referenced models of working memory. In this model, working memory is conceptualized as a system of three interacting parts: the phonological loop, visuospatial sketchpad, and central executive.

The phonological loop functions to retain verbal or acoustic information for a short period of time (Baddeley & Logie, 1999; Swanson, 2013). Much of this system comprises language processing and short-term memory span. Although a major component of working memory, short-term memory is not enough to account for all the variation in working memory, suggesting that although the systems are related, they are also distinct (Swanson, 2008). The visuospatial sketchpad is utilized to process and store

visually presented information in the mind's eye, and for recoding verbal information into an iconic form (Baddeley, 2007; Swanson, 2013).

The central executive system is considered to be the dominant system in the working memory model, as Baddeley and Hitch (1974) conceptualized the other two components to be “slave systems.” A common element of the central executive that relates to working memory is attentional control (Engle et al., 1999; Swanson, 2008; Unsworth & Spillers, 2010). Taken together, the central executive (as measured by tasks of attentional control) and short-term memory account for a majority of the variation in working memory scores (Swanson, 2008).

Later, Baddeley (2000) introduced the concept of the Episodic Buffer to his model to account for the interactions between working memory and long-term memory. The episodic buffer serves to both recall information from long-term memory into working memory in order to integrate information effectively with previously learned knowledge, and to store new information in long-term memory. Swanson (2008) indicated a strong relationship is present between naming speed and short-term memory. In addition, the relationship between speeded naming is also directly associated with performance on working memory measures (Bayliss, Jarrold, Baddeley, Gunn, & Leigh, 2005; McCabe & Hartman 2003). Based on the operational definition of the episodic buffer, these two tasks may be interrelated, as Naming Speed (Glr-NA) has been identified in Cattell-Horn Carroll (CHC) models of intelligence as a narrow ability under the factor of long-term retrieval (Flanagan, Alfonso, & Mascolo, 2006; McGrew &

Wendling, 2010), and long term memory in general has been identified as a contributor to choice of strategy when encoding information into working memory (Baddeley, 2012).

### **Statement of the Problem**

Dynamic assessment is not without its problems. Several challenges to the validity, reliability, and usefulness of this procedure have been made, including vague constructs, psychometrically inadequate instruments, and issues pertaining to the cost and time of assessment, as well as participants most likely to benefit from its application (Caffery, Fuchs, & Fuchs, 2008; Grigorenko & Sternberg, 1998; Jitendra & Kameenui, 1993; Murphy & Maree, 2006; Swanson, 1996). Age is one variable that is suitable for such exploration since several questions emerge related as to whether the effectiveness of dynamic assessment only has relevance to certain age groups. For example, there are some studies that suggest age is a relevant factor to consider when using dynamic assessment. Swanson and Lussier (2001) identified age as one of the variables of interest in their meta-analysis, and a main effect was noted. Specifically, these authors indicated greater effect sizes were present for children under the age of ten years, and no significant difference was found between children 10-13 years old and older than 13. Although it is meaningful, this finding is not without its limitations. Because this difference was based on a meta-analysis, the authors were unable to account for differences in experimental methodology. From this, one may ask if the modifiability of working memory, via dynamic assessment, and its prediction of academic performance is only applicable to younger age groups of children? To date, there have been no experimental studies that have systematically evaluated the effect of age on dynamic assessment.

## **Visual and Verbal Working Memory**

In addition to age, modality of working memory is also an important factor to consider. When using static assessment of working memory, some research suggests visual working memory predicts variation in achievement differently than reading. For example, verbal working memory has been considered the superior predictor of reading achievement over visual working memory tasks (Engle, Cantor, & Carullo, 1992; Swanson, 1992; Swanson, Zheng, & Jerman, 2009), while visual spatial working memory appears to hold a closer relationship to mathematics (Caviola, Mammarella, Lucangeli, & Cornoldi, 2014; Holmes et al., 2008). Wilson and Swanson (2001) found that both modalities were able to predict math performance; however, verbal working memory accounted for the majority of the variation in scores.

Strong evidence has been presented that suggests working memory is domain general, meaning the system works together in unison regardless of the modality of processing used to encode information. Swanson (2003) conducted a cross-sectional study with participants ages 7, 10, 13, and 20 years. He found differences in working memory scores between two groups of readers (learning disabled and average) became larger relative to improvements in phonological and visual performance, and this relationship varied based on the age of participant. Because processing efficiency measures appeared relatively equal between the two groups of readers, Swanson (2017) indicated memory span/storage was more likely to moderate working memory performance as capacity demands were greater for the learning disabled group. In addition, processing modality (visual/verbal) did not play a role in this finding, as there

was no noted interaction effect between them. It was also noted that capacity demands were similar for participants ages 7, 10, and 13 years, but were significantly less than the demands placed on 20 year olds. Swanson found that differences between learning disabled and non-learning disabled groups were present in the initial scores (static assessment), but the magnitude of these differences were greater for dynamic assessment measures (gain and maintenance scores). In addition, working memory performance in learning disabled participants was found to be domain general. In other words, no difference between performance on visual and verbal working memory tasks was present.

Although both verbal and visual working memory tasks may share a common factor (e.g., second order factor), it may also be the case that unique variance related to verbal and visual-spatial working memory play an important role in predictions of specific academic domains such as reading or math. More recently, when focusing on just initial scores of working memory, Swanson (2017) tested whether verbal and visual working memory tasks work as a singular system when predicting achievement. When evaluating performance in a large sample of individuals ages 6, 7, 8, 9, 11, 13, 15, 18, 26, 41, and 66, he found that age-related performance on working memory tasks appears to be best described as a single, general factor with only one exception occurring at the age of 18 years. When applying working memory performance as a general factor to prediction of achievement, hierarchical linear modeling indicated a general construct appears to underline age-related performance in achievement, likely because of the subordinate role each of these systems play to the central executive (Swanson, 2017).

However, it is important to note he did not determine whether working memory operated as a domain specific factor under dynamic testing conditions, a focus of this study.

Thus, further analysis is needed to explore the unique contribution of dynamic assessment of visual and verbal working memory to achievement measures. In addition, the extent to which age interacts with the relationship of this variable and its ability to predict unique variance in reading and math beyond that of the initial scores is unknown.

### **Purpose of the Current Study**

The current study seeks to address these gaps in the literature by evaluating the impact age has on the potential for dynamic assessment of working memory to predict the academic, domain-specific outcomes of reading and math. In the current study, working memory tests were administered with scaffolding support in order to improve the examinee's performance and obtain the highest level of memory span performance with assistance. A core assumption of these tests is that the probes, or series of cues presented to remind the participants of the information in each query, assist in the consolidation of memory and/or the ability to retrieve information from storage. The gain score, which is optimal performance under probe conditions, serves as a measure of maximized processing efficiency and maximizes working memory performance, bringing one's score closer to their optimal ability. The number of probes administered corresponds to processing efficiency, as the ability to progress through tasks with fewer hints suggests a higher degree of efficiency, and the ability to sustain optimal performance demonstrated with assistance given in the form of probes after a delay and without additional

assistance. This final probe is calibrated to align more closely to the examinee's optimal level of performance on a working memory task.

These measures utilize the following logic (also see Swanson 1999; 2003; 2011). The initial, or static condition reflects baseline performance on a working memory task without assistance from the examiner. The cueing conditions improve the examinee's ability to access stored information because probes are set to assist participants in retrieving information or in forming memory traces that were not thoroughly established under the baseline condition. Previous research has indicated such cues can result in higher performance, with some gaining as much as one standard deviation from their initial or baseline performance (Swanson, 1992; 1993). Because sequential processing is emphasized during probe presentation, the use of other, less effective recall strategies is de-emphasized. The re-administration of the highest item achieved with cues (referred to as the maintenance score) utilizes the same test items, but is administered without the support. By matching the maintenance item to the gain score item, the examiner is able to identify the differences in processing between age groups beyond the ability to learn the task. Since each of the participants were given items corresponding to their optimal level of performance, a decrement in the outcome when compared to gain scores is related to difficulties in processing capacity.

Given this overview of the literature, three research questions are proposed.

1. Does the dynamic assessment (as defined by measures of gain, maintenance) of working memory contribute unique variance in predictions of reading and math

achievement beyond the contribution of static or initial score (working memory performance without dynamic assessment) performance?

2. Are the effects of dynamic assessment (gain scores, maintenance scores) of working memory on reading and math achievement outcomes moderated by age?
3. Are the effects of dynamic assessment of working memory on reading and math achievement outcomes moderated by the interaction between the modality type of dynamic assessment (visual, verbal) measure and age?

The null hypothesis for question one states there is no difference between the capacity for dynamic and static assessments of working memory to predict reading or math outcomes ( $H_0: \textit{Static} = \textit{Gain} = \textit{Maintenance scores}$ ). The alternative hypothesis states that such differences are present, with dynamic assessment predicting a greater amount of variation in both reading and math ( $H_A: \textit{Static} \neq \textit{Gain or Maintenance}$ ). It is expected that in line with previous research, the null hypothesis to this question will be rejected in favor of the alternative.

For the second question, the null hypothesis being tested states there are no interaction effects present between age and dynamic assessment of a domain general construct of working memory when predicting reading or math ( $H_0: \textit{Age} \times \textit{DA (Domain General)} = 0$ ). Conversely, the alternative hypothesis states that interactions exist ( $H_A: \textit{Age} \times \textit{DA (Domain General)} \neq 0$ ). It is expected that interaction effects will be found.

The null hypothesis for the third research question states no interaction effects are present between age and dynamic measures of working memory when the construct of



working memory is bifurcated into verbal and visual modalities

( $H_0: \text{Age} \times \text{DA} (\text{specific modality}) = 0$ ). The alternative hypothesis states interactions are present ( $H_A: \text{Age} \times \text{DA} (\text{specific modality}) \neq 0$ ). It is expected that age by dynamic assessment interactions will be present when analyzing domain specific modalities.

## Chapter 2: Literature Review

Dynamic assessment is a measurement procedure that seeks to evaluate the modifiability of an individual. Lev Vygotsky (1978) is most often described as a primary source from which dynamic assessment emerged (Campione & Brown, 1987; Grigorenko & Sternberg, 1998; Haywood & Tzuriel 2002; Murphy & Maree, 2006). Vygotsky identified two zones of performance, which include the zone of actual development, and the zone of proximal development (ZPD). The first of these has been defined as the level of performance an individual can reach independently. This level of performance is often described as static. The ZPD references the difference between what a person does independently, and what can be done when provided with assistance/support by a more knowledgeable individual (Guthke & Wingefeld, 1992; Haywood & Tzuriel, 2002). Vygotsky held that it was within the ZPD that maturation and development occur. In addition, this development occurs as the synthesis resulting from the interaction between the teacher and learner; therefore, assessment within an interactive context is necessary in order to measure one's aptitude or proximal development (Murphy & Maree, 2006).

Although several protocols for assessment of the ZPD have been developed, Vygotsky never empirically validated this theory, as he died before such work could occur (Tudge & Scrimsher, 2003). In addition, he did not provide guidance as to how the ZPD should be measured (Guthke & Wingefeld, 1992; Tudge & Scrimsher, 2003). Fortunately, this theory was put into practice under the title of dynamic assessment by several of Vygotsky's successors, and different models to attempt to explain functioning within the ZPD have been proposed (See Grigorenko & Sternberg, 1998 for a

comprehensive review). These include Feuerstein's model of cognitive modifiability (Feuerstein, Rand, & Hoffman, 1979), Budoff's learning potential model (Budoff, 1987), Campione and Brown's transfer/graduated prompting model (Campione & Brown, 1987), Carlson and Wiedl's testing the limits model (Carlson & Wiedl, 1992), Guthke's learning test model (Guthke & Wingenfeld, 1992), and Swanson's Cognitive Processing Test (Swanson, 1992; 1996).

This review will provide brief explanations of the six major models of dynamic assessment. This will be followed by a review of the weaknesses of dynamic assessment and how they may be addressed. Special attention will be given to Swanson's (1992) psychometric application of dynamic assessment of working memory, and how this method is able to address many of the major criticisms of dynamic assessment, with cost and time remaining as potential challenges. This review will conclude with a summary of what is known about age as a moderator of dynamic assessment, followed by the research questions and hypotheses that guide the current study.

### **Feuerstein – The LPAD**

An Israeli psychologist named Reuven Feuerstein was one of the original leaders in the field of dynamic assessment. He developed the Learning Potential Assessment Device (LPAD) as a culturally unbiased tool to measure learning potential (Feuerstein, Rand, & Hoffman, 1979). Feuerstein believed assessment and feedback/support ought to occur at the same time (Grigorenko & Sternberg, 1998). He and his colleagues believed the LPAD differed from traditional assessments of intelligence as the subtests are meant to concurrently teach and test an ability rather than only testing it. Other differences

between the LPAD and traditional assessment included an emphasis on the process one uses to derive an answer over the score itself, the way findings are interpreted, and the interactive nature of the assessment process. Feuerstein and colleagues believed that because dynamic assessment results indicate interventions and strategies that can be useful in remediation, this process would rise above traditional, static testing, which he believed to be a pessimistic process (Grigorenko & Sternberg, 1998). Thus, he believed dynamic assessment would ultimately replace static, traditional evaluations (Feuerstein et al., 1979). Unlike Vygotsky, who emphasized social interaction within the context of assessment, Feuerstein emphasized modification of the child's cognitive ability (Murphy & Maree, 2006).

Feuerstein's model (Feuerstein et al., 1979) suggests assessment requires modification in four areas: the structure of the test, the relationship between the examiner and examinee, shifting the focus of the assessment towards a process-based approach, and methods of interpretation through the analysis of peaks and valleys rather than the standard score. Impairments were modelled based on a three-part conceptualization of the test: the input, elaboration, and output of response, and the interactions between these phases were important in knowing the extent and significance of the impairment. The input phase focused on the breadth and depth of information the individual was able to use to encode a problem. Processes that relate to input include impulsive or poor spatial concepts. Elaboration referred to the individual's ability to draw on information for processing. Examples of problems that affect this area include poor selective/focused attention, under-developed perception of the problem, and a lack of what Feuerstein

called spontaneous comparative behavior. Outputs, or the method in which the child relays the response, can be affected by a number of factors including impaired verbal ability, trial and error responding, and impulsivity. Both emotional and motivational processes were also considered to be important moderators of performance throughout the process.

For Feuerstein and colleagues (1979), dynamic assessment held five goals. First, the extent to which an individual's performance can be changed is measured through presentation of assessment materials that create structural changes. Second, it is important to measure both the limitations and conditions of the environment in which modification can occur. The third goal of assessment is to identify the amount of assistance necessary to elicit change in the individual, while the fourth goal is to assess the importance or significance of any modification in the person's performance that is demonstrated, particularly in transfer to untrained conditions. The fifth goal of assessment is to note and record the individual's preferred modalities of processing and both strengths and weaknesses.

### **Budoff – The Training Based Method**

The training based method (Budoff, 1987) of dynamic assessment was constructed under the premise that intelligence and trainability are closely related. In this model, training is given on items that are independent to both the pretest and posttest but are directly related to the task being assessed, and posttests are used to evaluate the degree of change within the individual post training. Budoff's model follows the test-teach-test model and is centered on modification of cognitive structures.

The purpose of Budoff's method was to distinguish between levels of severity in populations with intellectual impairments, and to improve classification of those with more mild cognitive impairment. After receiving interventions, individuals in his studies were classified into three groups. *High Scorers* consisting of those who did well on the assessment without training, *Gainers*, who did poor initially but showed growth post intervention, and *Non-Gainers* who did not show benefit from instruction (Grigorenko & Sternberg, 1998). In assessing these groups, Budoff relied on all three types of scores he attained (the pretest score, posttest score, and the residual gain score) with standardized testing administration being used for both pretesting and post-testing conditions. Using posttest scores, he found that trained students profited from a systematic learning experience more than the non-trained controls, indicating support for the predictive validity of his model (Budoff, 1987).

Budoff's approach differed from his predecessors because it was purposefully designed to improve classification of children into special education rather than identifying strengths and weaknesses. In addition, his approach relied on psychometrically valid instruments as part of the assessment process, and the teaching component of his model targeted student understanding of the task demands (Grigorenko & Sternberg, 1998). Consistent with several methods of dynamic assessment, Budoff was resistant to the use of academic measures in establishing criterion validity, which arguably represents one weakness to his methods (Grigorenko & Sternberg, 1998) when applied within educational psychology.

### **Campione and Brown – The Graduated Prompting and Transfer Model**

Campione and Brown (1987) introduced the graduated prompting and transfer model of dynamic assessment. Their method is guided by information processing theory and was used to target students who were academically underperforming (Grigorenko & Sternberg, 1998). In this model, the child is given a series of least-to-most prompts, or progressively revealing hints to assist in solving the problem. These hints are created based on a task analysis, and progress from more general observations to specific forms of help. This method is helpful when one wants to identify the minimum amount of assistance the child needs in order to complete a task, and it is helpful for establishing what types of support are most effective for the child. This model avoids the psychometric problems associated with change scores because the metric of measurement becomes the amount of assistance given, rather than the child's score. In addition, tests that give partial credit scores based on the number of prompts given have been found to be more sensitive to change in individuals with intellectual disabilities (Haywood & Tzuriel 2002; Tzuriel & Klien, 1985). Transfer, or the ability to utilize the help that is given and apply it to future problems, is also identified as an important element by Campione and Brown (1987).

Graduated prompting is task-focused rather than child-focused, which distinguishes it from other forms of dynamic assessment. In addition, neither standardization of interventions nor transfer were used in dynamic testing prior to Campione and Brown, which in effect changed the focus of assessment away from the examinee and towards the assessment itself (Grigorenko & Sternberg, 1998).

### **Carlson and Wiedl – Testing the limits Model**

Carlson and Wiedl (1992) developed the testing the limits model, which was designed to improve the testing performance of disadvantaged youth. In this model, the change score was avoided, as modification within testing was a focus. The child is administered static, traditional assessments with adherence to the standardization process. Carlson and Wiedl (1992) regard this score as the final score. After standardized administration is concluded, assessments are administered again, but with modifications including prompting procedures and provision of feedback that ranges from simple to elaborate. This allows the assessor to evaluate how non-cognitive variables impact testing performance, and to determine which variables lead to improved performance (Grigorenko & Sternberg, 1998).

To accomplish this, Carlson and Wiedl (1992) developed three levels of assessment analysis. The first of these include personal factors such as skills that comprise an individual's cognitive performance and are used to create procedures for solving a problem. Examples of such factors may include a mix of cognitive abilities such as executive functions and visual scanning. Next the task requirements are analyzed and the interaction between the personal factors and the task are observed. Diagnostic approaches comprise the third component and focus on long-term metacognitive modification, as well as short-term compensational strategies.

Carlson and Wiedl (1992) made many important contributions to the field of dynamic assessment. Their use of standardized assessments as both a pretest and posttest made their approach uniquely different from Feuerstein's (1979). In addition, they



introduced personality as a variable in dynamic assessment by collecting data on examinee characteristics such as introversion, impulsivity, and neuroticism, and introducing them into their models (Carlson & Wiedl, 1979; Grigorenko & Sternberg, 1998). Grigorenko and Sternberg (1998) noted that the Carlson and Wiedl (1992) method is based on group comparisons of performance; therefore, intra-individual differences are not evaluated, as studies relied on the assessment of group differences.

### **Guthke - The Learning Test**

Guthke and Wingefeld (1992) hypothesized that measuring learning that occurs during an assessment should help examiners make a more accurate diagnosis. In addition, they believed dynamic assessment could make more valid predictions about the trajectory of cognitive growth and it could lead to the development of better recommendations for intervention. Their methods were similar to Vygotsky's theory, as they argued for the necessity of measurement of standardized (or static) ability as well as functioning within the ZPD. They offer assessment variation within their learning potential battery by providing a long-term battery, which includes a training process that takes several days, and a short-term battery that was developed to measure modifiability within the assessment. In describing the learning test model, Grigorenko and Sternberg (1998) stated that Guthke and colleagues utilized a number of methods such as systematic feedback, prompts, and repetitions, although these processes are a larger part of the long-term test, whereas the short-term test incorporated detailed and simple feedback into the assessment process.

Guthke and Wingenfeld (1992) recommended using the posttest scores rather than dealing with the psychometric problems associated with change scores as all learning accrued by the individual was reflected in this score. Thus, the posttest represents the ability of the student by marking his or her position within a normal distribution after intervention was delivered and received. Grigorenko and Sternberg (1998) cite multiple studies by Guthke and colleagues that support the findings that the posttest score within the learn test model is more informative than the training or baseline scores.

Guthke and Wingenfeld (1992) also focused their work on differentiating severity levels of intellectual disabilities. According to their work, individuals with intellectual disability showed improvements on posttest intelligence scores, and unlike the pretest scores, the measures given after training were able to predict future performance. They found that conducting a task analysis of assessments was needed in order to focus tasks into a process-oriented assessment, and that the best way to accomplish this is when using tests that are domain-specific rather than tests focusing on general intellectual ability. They did not believe standardized mediation was necessary, as such a process would interfere with the greater need to develop individualized training for the purpose of discovering how the child works through problems. Thus the nature of mediation varied depending on the individual responses and needs of the individual.

### **Swanson – The Cognitive Processing Test**

Swanson (1992; 1995; 1996) developed the Swanson – Cognitive Processing Test (S-CPT) in order to provide a standardized index for measuring processing potential and an individual's declarative knowledge of planning to remember, as well as to assist in the

identification of strengths and weaknesses of an individual. This assessment was later update, and its name was changed to the abbreviate Test of Working Memory (aTWM; Swanson, 2013). The aTWM included updates to the psychometric properties of the assessment, as well as an elimination of many of the original subtests in order to shorten administration time.

An assumption of the S-CPT/aTWM is that individual differences in working memory account for variation within primary processing modalities (Swanson, 1996). Swanson suggested the majority of all information processing models contain components of working memory, and that this skill is highly correlated with achievement, particularly reading comprehension and mathematics. Such a relation between working memory and achievement is well documented within the literature (Gathercole, Lamont, & Alloway, 2006; Swanson 2013). Previous research has demonstrated a relation between verbal working memory and reading (Daneman & Hannon, 2001; Swanson, Zheng, & Jerman, 2009), and visual spatial working memory and math (Caviola, Mammarella, Lucangeli, Cornoldi, 2014; Holmes, Adams, & Hamilton, 2008). Wilson and Swanson (2001) noted that both verbal and visual working memory contributed to performance on a measure of mathematics, although the majority of the variation was explained by verbal working memory.

Additionally, these findings have also been confirmed by meta-analyses of academics and Cattell-Horn Carroll (CHC) theory (Flanagan, Alfonso, & Mascolo, 2006; McGrew and Wendling, 2010) as well as Swanson's research (Swanson, 2008; 2017).

Due to this relation, working memory appears an excellent choice of construct for applying dynamic assessment.

Multiple studies analyzed the construct and criterion-related validity of the S-CPT/aTWM as well as its reliability (Swanson, 1992; 1995; 1996; 2013). Results of these studies suggest that these assessments are better able to predict achievement in the areas of reading and math than the WISC-R, and that the gain score was the only variable able to determine individual differences in mathematics performance. He also found that children with true learning disabilities demonstrated stability between the pretest and posttest conditions, supporting the hypothesis that children with true learning disabilities demonstrate low levels of modifiability. In conclusion, Swanson indicated that his studies on the dynamic assessment of working memory support the test's ability to provide educators with instructionally-relevant information on (1) how effective simple feedback is, (2) the child's knowledge of strategies to solve problems, (3) the extent to which training is generalized once probing procedures have been removed, (4) how flexible the child is to intervention, and (5) the child's preference between visuospatial and verbal modalities (Swanson, 1996).

More recently, Swanson (2010) conducted a three-year longitudinal study, evaluating the potential for dynamic assessment of working memory to predict growth in phonological awareness and vocabulary. In children with and without reading disabilities, initial, gain, and maintenance scores from two subtests taken from the S-CPT were administered three times at one year intervals. Results of the study indicated assessments of working memory under the initial and maintenance conditions predicted growth in

receptive vocabulary, while nonword fluency growth was predicted by the gain condition. Thus, Swanson concluded that the dynamic assessment of working memory was related to growth rates of both vocabulary and decoding.

Swanson (2011) continued to evaluate the effectiveness of dynamic working memory measures by evaluating its ability to predict growth in reading comprehension in a sample of children with learning disabilities. Four measures (two verbal and two visual) of working memory were used, with the two verbal measures (*Semantic Association* and *Digit/Sentence Task*) being the same as those used to predict decoding and vocabulary Swanson's earlier work (Swanson, 2010). Results of the study indicate the working memory maintenance score derived during the first year of the study was able to predict an additional 40% of the variation in reading comprehension scores collected during year three. After controlling for the scores at initial testing, maintenance scores for verbal memory explained an additional 7% of variation in year three scores, and this was supported through growth modeling procedures. Taken together, these two studies provide evidence of predictive validity for dynamic assessment of working memory in determining future academic performance.

### **Age related factors and working memory**

Swanson (2017) noted that when assessed using static measures, the trajectory of working memory across different age spans is not clear. There are several studies that suggest working memory develops in a linear trend from ages four to ten years, with a flat trend continuing from age fifteen up (Gathercole, Pickering, Ambridge, & Wearing, 2004; Luciana, Conklin, Hooper, & Yarger, 2005; McAuley & White, 2011); however,

development up to age twenty has also been noted (Hamilton, Coates, & Heffernan, 2003) and there were even changes noted around the age of thirty in adulthood (Alloway & Alloway, 2013).

Nevertheless, several components of working memory are known to play important roles in the development of this system across age groups. For example, storage of information has been shown to predict age related changes in working memory (Bayliss, et al., 2005; Swanson, 2008). Swanson (2008) found that children ages six to seven and ages eight to nine did not differ in regard to how memory was structured, but did note that controlled attention (particularly fluency and random generation) predicts age-related differences in working memory, while speed and phonological awareness predicts age-related differences in short term memory or the phonological loop.

How working memory tasks are generally approached across the lifespan also suggests that age-related differences in its development should be present. di Ribaupierre, Lecerf, and Bailleux (2000, as cited by Swanson, 2017) found that adults are more likely to use verbal strategies for both visual and verbal tasks, while children are more likely to use visual-based strategies through the age of ten years. Gathercole (1998) found memory strategies to be unstable in younger children, which suggests differences in how a working memory task is approached depending on one's age. Swanson (2008) indicated procedural strategies related to working memory tend to become more stable at eight and nine years of age versus the unstable time period of six to seven years. In addition, he indicated the ability to process language is a slower task for young children when

compared to that of older children, and this reduction in speed is related to poor performance on tests of working memory

Swanson (2017) examined age-related differences in visual and verbal forms of working memory. He sought to evaluate whether visual and verbal working memory formed a single construct of working memory or could be better understood as independent processes. He evaluated the evidence for a domain general versus domain specific system of working memory across the age span.

Less information is known about the impact age has on the dynamic assessment of working memory. The study that comes closest to evaluating the impact of age on dynamic assessment was the meta-analysis conducted by Swanson and Lussier (2001), which included age as one of the variables of interest in this study, and a main effect was noted. Specifically, Swanson and Lussier indicated greater effect sizes were noted for children under the age of 10 years, and no significant difference was found between children 10-13 years old and older than 13.

Although it is meaningful, this finding is not without its limitations. Because this difference was based on a meta-analysis, the authors were unable to account for differences in experimental methodology. In fact, it was noted that these studies varied widely in both their outcomes and their applied methodologies, and it was noted that multiple studies included in the analysis used dynamic assessment procedures that were not psychometrically adequate. Additionally, the majority of the articles evaluated the modifiability of intelligence scores rather than achievement.

In addition to the effects of age in moderating the variance found in achievement scores, additional research is needed to explore the differences in processing modalities within dynamic assessment of working memory. Within the literature, there is evidence that static, traditional assessment of working memory may vary its predictive power of achievement scores, when comparing math versus reading as the outcome, and visual versus verbal measures of working memory as the independent variable. For example, verbal working memory has been considered the superior predictor of reading achievement over visual working memory tasks (Engle, Cantor, & Carullo, 1992; Swanson, 1992; Swanson, Zheng, & Jerman, 2009), while visual spatial working memory appears to hold a closer relationship to mathematics (Caviola, Mammarella, Lucangeli, & Cornoldi, 2014; Holmes, Adams, & Hamilton, 2008). Wilson and Swanson (2001) found that both modalities were able to predict math performance; however, verbal working memory accounted for the majority of the variation in scores. When evaluating the impact of a domain general system of working memory on achievement, Swanson (2017) found evidence for a singular construct of working memory. Whether such findings hold under dynamic assessment conditions is currently unknown, and the extent to which age is able to moderate the prediction of achievement within each type of modality and across domains is also unknown.

### **Purpose of the Study**

The current study seeks to determine whether dynamic assessment measures contribute unique variance to achievement measures, but more importantly to determine



the degree to which potential moderating variables, such as age, interact with dynamic assessment predictions of achievement. Three research questions are proposed.

1. Does the dynamic assessment (as defined by measures of gain, maintenance) of working memory contribute unique variance in predictions of reading and math achievement beyond the contribution of static or initial score (working memory performance without dynamic assessment) performance?
2. Are the effects of dynamic assessment (gain scores, maintenance scores) of working memory on reading and math achievement outcomes moderated by age?
3. Are the effects of dynamic assessment of working memory on reading and math achievement outcomes moderated by the interaction between the modality type of dynamic assessment (visual, verbal) measure and age?

The null hypothesis for question one states there is no difference between the capacity for dynamic and static assessments of working memory to predict reading or math outcomes ( $H_0: \textit{Static} = \textit{Gain} = \textit{Maintenance scores}$ ). The alternative hypothesis states that such differences are present, with dynamic assessment predicting a greater amount of variation in both reading and math ( $H_A: \textit{Static} \neq \textit{Gain or Maintenance}$ ). It is expected that in line with previous research, the null hypothesis to this question will be rejected in favor of the alternative.

For the second question, the null hypothesis being tested states there are no interaction effects present between age and dynamic assessment of a domain general construct of working memory when predicting reading or math ( $H_0: \textit{Age} \times \textit{DA (Domain General)} = 0$ ). Conversely, the alternative hypothesis states

that interactions exist ( $H_A: Age \times DA (Domain\ General) \neq 0$ ). It is expected that interaction effects will be found.

The null hypothesis for the third research question states no interaction effects are present between age and dynamic measures of working memory when the construct of working memory is bifurcated into verbal and visual modalities ( $H_0: Age \times DA (specific\ modality) = 0$ ). The alternative hypothesis states interactions are present ( $H_A: Age \times DA (specific\ modality) \neq 0$ ). It is expected that age by dynamic assessment interactions will be present when analyzing domain specific modalities.

I hypothesize that dynamic assessment of working memory will predict unique variance above that of the initial scores, thus replicating earlier findings (e.g., Swanson, 1992). I also hypothesize that different modalities of dynamic assessment (visual versus verbal) will predict variation in reading and math scores differently, with visual having a stronger relationship with math scores and verbal measures aligning more closely to reading scores. Finally, I predict that the effectiveness of dynamic assessment in predicting variation in achievement will be moderated by age, with the younger ages showing the most benefit.

## Chapter 3 - Methods

### Participants

A total of 602 participants from a secondary data set were included in the study. These data were collected between 1990 and 2012 with the earliest participants in the study being the same individuals who participated in the Swanson (1992) study. Subsequent data were taken from additional studies of dynamic assessment and were then edited and combined together for the current analysis. Testers were graduate students from primarily five universities (University of California, University of British Columbia, University of Washington, University of Miami, and the University of Northern Colorado, see Swanson, 2013, for details). All examiners were trained to administer standardized assessments, with specific instruction given in administering and scoring the dependent and independent measures.

The current sample represented a diverse group of individuals with a mean age of 12.25 years ( $SD=4.20$ ) and included 358 boys and 340 girls. Of these participants, 471 students were identified as white, 41 as black, 58 were Latino, 16 were Asian, and 18 identified as "other." Data pertaining to ethnicity were either missing or not completed by the participants in 94 cases. Information regarding the socioeconomic status of the sample was not complete and was thus not included. However, according to the testing manual (Swanson, 2013), the sample included a broad range social economic groups.

A one way ANOVA found no significant differences between males and females on either the reading,  $F(1, 323)=.518, p = .47$ ) or mathematics,  $F(1, 322)=1.29, p=.26$ ) measures. Differences in performance were, however, noted between ethnicity groups on

the math  $F(4, 319)=7.21, p<.001$ ) and reading  $F(4, 320)=14.08, p<.001$ ) measures. Interpretation of these should be made with caution, as non-white participants were significantly under- represented in this sample.

Because an interaction effect was noted in the regression models between dynamic assessment scores and age, individuals in the study were grouped by age for a follow-up analysis, and the groups were created to reflect one's position in K-12 education. These subgroups consisted of individuals ages five through seven years (early elementary,  $n=80$ ), eight to nine years (middle elementary,  $n=115$ ), 10 to 11 years (late elementary,  $n=166$ ), 12 to 13 (middle school,  $n=117$ ), and 14 through 17 (high school,  $n=96$ ).

### **Dependent (Criterion) measures**

Two outcome measures from the Wide Range Achievement Test (WRAT, Jastak & Wilkinson, 1984; Wilkenson, 1993) were used to measure achievement. The reading subtest is a measure of word identification. On this measure, a list of words is presented to the examinee in increasing levels of difficulty. A response is considered to be correct if the word is read both accurately and fluidly. If a word is read correctly, but the examinee has to sound it out, the response is marked as wrong. Math performance was assessed using the Arithmetic Computation subtest. This assessment requires the examinee to solve a series of computations ranging in difficulty from simple addition to complex algebra. The WRAT-3 assessment manual reports a test-retest reliability coefficients ranging from .81 to .98 for the subtests. In addition, evidence of concurrent validity is

presented with a correlation of .40 to .70 noted between the WRAT -3 and the Woodcock Johnson test of achievement and Wechsler Individual Achievement Test.

### **Independent measures**

Four measures from Swanson's normative work in working memory (Swanson, 1996; 2013) were utilized as part of the study. These include the Semantic Association (referred to in recent studies as Conceptual Span), Auditory Digit Sequence, Visual Matrix, and Mapping/ Directions subtests. For each of these measures, a baseline, gain, and maintenance score was established. The baseline measure is consistent with traditional, static assessment, while the gain and maintenance scores represent the dynamic assessment testing conditions. Scores utilized for the current study consisted of a mean composite of z-scores, where the Concept Span/Semantic Association and Digit Sentence Task formed the verbal working memory composite score, and the Visual Matrix and Mapping/Directions scores formed the visual working memory composite score (Swanson, 2010; 2011). In other words, the baseline or initial working memory score z- scores from each of the four subtests, and the gain and maintenance z-scores were based on the mean and standard deviations of the initial scores, respectively. When analyzing the visual and verbal domain specific modalities separately, a similar process will be used in which the mean composite z-scores reflect performance on the two verbal measures to form a composite score, and on the two visual measures to form the visual working memory composite scores.

**Semantic Association/Conceptual Span.** The purpose of this task is to measure a child's ability to organize information presented verbally into abstract categories. Here,

the examiner provides the examinee with a set of words, asks a question that requires the examinee to process information relating to the set of words, and then asks the examinee to recall words that go together (Swanson, 1996, p. 66). This test has a total of eight sets of items that range from two semantic categories and four total words, to four categories and 16 words. The following directions are given to the examinee:

“I am going to say words. Some of the words go together. Don’t tell me the words in the order I give them to you, but say the words that go together. For example, if I say the words “car, baseball, truck, football,” you would say “car and Truck” first because they go together, and then you would say “baseball and football” because they go together”... (Swanson, 1996, p. 66)

Further elaborations are provided, and a second example is also offered if the examiner believes the child is struggling to understand the task. The score from this assessment is the number of the last item set that was recalled correctly, and possible raw scores range from zero to eight.

***Gain score.*** For each incorrect response, a series of probes is given. Probes for this test are as follows: First, the examiner informs the examinee of all the category names and the final word that appears in the list within each category. Second, the category names are repeated and the first word from each category that appears in the list is given. The third probe involves listing the category with all words in the medial position, while the fourth probe is a re-administration of all the words in the original order. Probes are given until the examinee is able to repeat all of the words in the correct sequence, and the assessment is discontinued if he or she is unable to provide a correct

response after all four probes are presented. The gain score then represents the highest numbered set that is recalled correctly with probes.

**Maintenance.** After a delay, the participant is asked to repeat the words again. The examiner presents the list of words that represented the examinee's highest level of performance when probes were present; however, this time the assessment is given without assistance. The examiner says "The set of words that I'm going to read to you was presented earlier. I want to see if the set is now easier for you to remember (Swanson, 1996, p. 67)." If the participant makes an error or is unable to recall the words in the correct order, the maintenance score is the same as the initial score. If the examinee is able to recall the words in the correct order, then the maintenance score is the same as the gain score.

Swanson (1996) reported coefficient alpha scores of .80, .86, and .83 for the initial, gain, and maintenance scores respectively.

**Auditory Digit Sequence(Digit-Sequence Span).** This assessment measures the participant's ability to recall numerical information that is embedded within a short sentence. The numerical information references either a location or address. On this test, the examiner reads a sentence and then asks the examinee a process question. The examinee is then asked if he or she can recall the numbers in the sentence after explaining the strategy that will be used to recall the information. To begin the test, the examiner says,

"I'm going to read you some sentences that have information I want you to remember. All the sentences have to do with remembering an address, but I would

like you to pay attention to all the information in the sentence, because I will ask you a question about the sentence. After I present the information and before you recall it, I will ask you to choose a strategy, a way of remembering the information that you think would best help you remember.” (Swanson, 1996, p. 38)

After the information is given, the examiner then presents a strategy card that has pictorial representation of four different strategies one might use to help remember numbers. With this visual, a verbal explanation of each strategy is provided. As an example, the examiner may read a sentence that includes a street number and street name. The process question may be to ask the examinee what the name of the street was. After a response is given, the examinee is then asked to point to the picture on the strategy card that they will use to help remember the address. Finally, the examiner provides the prompt for the examinee to recall the numbers of the address in order. The initial score represents the last item set that is recalled correctly and independently.

***Gain Score.*** In the event the examinee makes an error, a series of up to four probes are given to assist him or her in reaching the correct response. These probes begin with provision of the last number, then move on to provision of the first numbers, the middle numbers, and then finally the numbers in order from the first to the last with the examinee being asked to repeat them. The gain score represents the highest numbered item that is recalled correctly under probe conditions.

***Maintenance Score.*** After a delay, the examiner returns to the Auditory Digit Sequence test and presents the examinee with the highest item that was recalled correctly



with probes. The examiner then says, "The sentence I'm going to read to you was presented earlier. I want to see if the numbers in the sentence are now easier for you to remember" (Swanson, 1996, p. 41). However, additional probe assistance is not given. If the examinee recalls the numbers correctly, then the maintenance score is equal to the gain score. If an error is made, then the maintenance score is equal to the initial or baseline score.

Coefficient alpha scores for the Auditory Digit Sequence subtest were .73, .82, and .81 for the initial, gain, and maintenance scores respectively (Swanson, 1996).

**Visual Matrix.** The purpose of this assessment is to measure an individual's ability to remember visual information that is presented sequentially and within a matrix (Swanson, 1992). Here, the examinee is given a series of dots within a matrix, as well as five seconds to review the matrix. The matrix is then removed from sight, and the examinee is asked a process question such as asking if there were any dots in the first column. After a response is given, the examinee is presented with a blank matrix and asked to draw the dots in the correct order (Swanson, 1996). To begin the test, the examiner says,

"I'm going to show you some pictures of boxes, and there will be dots in some of these boxes. I'm going to show these pictures to you for five seconds and then I will cover up the pictures. You have a sheet of paper with boxes in front of you. I would like you to fill-in where the dots are to go for each box. Before you begin filling in the dots, I will ask you a question." (Swanson, 1996, p. 30)

The initial score represents the number of the highest correctly produced matrix without any support provided by the examiner.

***Gain Score.*** When an error is noted, the participant is informed that he or she missed some of the dots. The examinee is then told that he or she will be given a series of hints. These hints consist of the examiner modeling correct placement of in predetermined columns (beginning, middle, end), and they continue in a least-to-most fashion, culminating with the examiner drawing all the dots for the examinee. The highest numbered matrix that is correctly recalled with assistance represents the gain score.

***Maintenance score.*** After the subtest has been completed and a short period of time has passed, the examiner then returns to the matrices and presents the examinee with the highest numbered matrix that was successfully completed with hints. This time, the matrix is presented with no help. The examinee is told, "This matrix that I'm going to show you was presented earlier. I want to see if the matrix is now easier for you to remember" (Swanson, 1996, p. 34). If a mistake is made on this item, the maintenance score is the same as the initial score; however, if it is correct, then the maintenance score is the same as the gain score.

Alpha coefficients for the Visual Matrix test, as reported by the S-CPT manual (Swanson, 1996) are .73, .79, and .76 for the initial, gain, and maintenance scores respectively.

**Mapping and Direction.** This subtest measures the examinee's ability to remember a visual spatial sequence of directions. The examinee is presented with a

"street map" and is then asked to review the map. After it is removed, a process question is asked in which the examinee must indicate a strategy that will be used to remember street and stoplights. The examinee is then asked to draw stoplights and street directions on a blank map. The following directions are read to the participant: "You can see on this map that there are buildings and streets. There are also dots, lines, and arrows. The dots are stoplights and the lines and arrows are directions. I want you to imagine you are driving a car and you are lost in the city. You asked for directions from people in the city and they drew you a map like this one. The map will help you get out of the city. Sometimes the car will zigzag on some streets, and that's okay, as long as you follow the arrows. Before you draw the directions and stoplights, I would like you to show me the way you are going to remember. I would like you to pick a picture that best. Matches how you plan to remember the directions and stoplights" (Swanson, 1996, p. 41, 43). An explanation of four strategies is then given along with pictorial cues. Participants are then presented with a map, and are asked a processing question after the map is removed. An example of such a question is, "Where there any stoplights in the first column" (p. 44). The examinee is then shown the picture of four strategies and is asked to identify the strategy he or she will use to remember. Then, the participant is asked to draw the lines, arrows, and dots as quickly as possible on a blank map. The initial score represents the highest numbered item that is completed correctly without assistance from the examiner.

***Gain score.*** When an error is noted, the examiner presents the examinee with probes that are meant to provide help with deriving an answer. These probes range from the examiner drawing in dots, lines, and arrows for specific columns, and proceed up to

the examiner providing all responses from of the item for the examinee. The gain score represents the highest numbered item that the individual is able to complete with help.

*Maintenance score.* After a break from this test, the examiner returns to the Mapping/ Directions subtest and informs the examinee, "This map that I'm going to show you was presented earlier. I want to see if the map is easier for you to remember this time" (p. 47). The highest numbered item that was completed with assistance (the gain score) is presented again, but without probes/help. If the examinee correctly responded to this item, the maintenance score is equal to the gain score. If an error occurs, the maintenance score is the same as the initial score. For this subtest, the S-CPT manual reports alpha coefficients of .72 for the initial score, .81 for the gain score, and .79 for the maintenance score (Swanson, 1996).

## Chapter 4 – Results

The research questions were evaluated by using two statistical procedures: regression and analysis of covariance (ANCOVA). A simultaneous regression model was used so that each variable would partial out the shared influence of the others within each step of the model. The purpose of the regression model was to assess the unique variance of dynamic assessment measures in predicting reading and math, specifically, the interactions between age and dynamic assessment. Regression was selected as the primary method of analysis because it holds several advantages in predictive models (Field & Miles, 2010). First, the evaluator is able to systematically enter variables into the model in order to account for the changes in its capacity to predict an outcome. Thus, changes in the predictive power of a model can be identified through systematic control of the variables of interest. In addition, regression analysis is able to provide beta values, which allow for analysis of the specific effect each independent variable will have on the dependent variable, and the degree of change in scores one would expect to see in a variable given a single unit change in the other.

Because age related differences in dynamic assessment was of concern in this dissertation, ANCOVA was used as a follow-up test to further explore interaction effects between age and dynamic scores. Age was a continuous variable in the regression modeling, whereas age was a categorical variable in the follow-up in the ANCOVA. The covariates in the ANCOVA were ethnicity and baseline (initial) scores.

## **Characteristics of the Data**

Table 1 includes descriptive information on the sample, including frequency counts for gender and ethnicity, and the mean, standard deviations, minimum, and maximum values for each of the dependent and independent variables. A one-way ANOVA indicated there was no significant differences between males and females on either the reading,  $F(1, 323)=.518, p=.47$ ) or mathematics,  $F(1, 322)=1.29, p=.26$ ) measures. Performance differences were, however, noted between ethnicity groups on the math  $F(4, 319)=7.21, p<.001$ ) and reading  $F(4, 320)=14.08, p<.001$ ) outcomes. Thus, interpretation of the subsequent outcomes must be interpreted with caution, as non-white participants are underrepresented in this sample. Thus, ethnicity was entered into each model of analysis in order to prevent it from serving as a possible intervening variable. This was accomplished by coding ethnicity into binary terms, with Caucasian/White participants falling in one group, and all other individuals into a second group.

## **Composite Scores and Correlations**

Rather than using a single test as a predictor, two or more measures were combined in order to create a more psychometrically robust variable. Scores utilized for the current study were developed by creating a mean composite of z-scores, where the Concept Span/Semantic Association and Digit Sentence Task formed the verbal working memory composites, and the Visual Matrix and Mapping and Directions scores formed the visual working memory composite (Swanson, 2010; 2011). In other words, the baseline working memory scores (i.e., initial scores) were the average of the sample initial z-scores from each of the four subtests, and the gain and maintenance z-scores

were the standardized average of the participant's performance on four gain scores and four maintenance scores based on the mean and standard deviations of the initial scores, respectively. When analyzing the visual and verbal modalities separately, a similar process was used in which the mean composite z-scores reflects performance on the two verbal measures to form a composite score, and on the two visual measures to form the visual working memory composite scores. In addition, the achievement data was converted to z-scores within age groups in order to simplify interpretation.

A correlation analysis was conducted in order to determine the relationship among the variables proposed for the models in this study, and results are displayed in Table 2. Results of this analysis indicated a moderate correlation between reading and each of the working memory composites, with verbal initial scores yielding the lowest correlation ( $r=.31$ ), and the domain general gain scores ( $r=.49$ ) yielding the highest. Similar correlations between the math and the independent variables were noted with verbal initial scores having the smallest correlation ( $r=.36$ ) and the domain general scores of gain and maintenance each having the largest correlation ( $r=.51$ ).

Potential collinearity was evaluated among the independent variables. Domain general initial scores yielded a strong correlation with both domain general gain ( $r=.80$ ) and maintenance ( $r=.84$ ) scores, while the latter two also strongly correlated to one another ( $r=.86$ ). Regarding the domain specific modalities, verbal and visual scores yielded moderate to high correlations for the initial ( $r=.58$ ), gain ( $r=.62$ ) and maintenance ( $r=.53$ ) conditions. In summary, correlations were considered to be moderate to high

across the variables within the data set, but were not considered to be high enough to suggest that any two variables were collinear (Cohen, 1988; Field & Miles, 2010).

### **Regression analysis**

A simultaneous regression model was computed using the domain general (model 1) and domain specific (model) working memory scores. Approximately half of the participants ( $n=304$ ) in the study were not administered the WRAT reading or math subtest, thus, they were not included in the regression analysis. This resulted in a total sample size of  $N=298$ . Prior to analyses, the assumptions of regression were tested to ensure the data were an appropriate fit for a linear model. This included checking for outliers, multicollinearity, homogeneity of variance, linearity, and normality.

To check for outliers,  $z$ - scores for each of the two dependent variables (reading and math) were computed. For the purpose of this paper, any value which fell three standard deviations from the mean or greater was considered to be an outlier (Field, 2015). No such values were noted within the reading scores; however, one individual did have a significantly discrepant  $z$ -score on the math outcome. Because only one individual was noted in an adjusted sample consisting of 298 participants, this datum was retained.

Homogeneity of variance was assessed by plotting the residuals against the independent variables. Results of this analysis indicated no significant concerns were present as the distribution of scores at each level appear to be equal. Regardless, any deviation in this assumption is not considered to be of concern as the results were analyzed using the method of least squares, which is able to produce unbiased results even when this assumption cannot be assumed (Field, 2015).



Independence was tested by plotting the standardized residuals against the predicted values of the model. Because no recognizable pattern was present, there is evidence to suggest the assumption of independence was satisfied. Additionally, the assumption of linearity was tested by plotting the residuals against each of the outcome variables. Visual analysis of these plots suggested no concerns with linearity.

Normality was tested by visual analysis of the outcome variables in a histogram and through analysis of a PP-plot. Results suggest the data follows an approximately normal distribution pattern. In addition, the Kolmogorov-Smirnov test was used to determine if the sample significantly differed from normal distribution. Results indicated no significant differences were noted on the reading outcome  $D(4, 293)=0.05, p=.12$ , but were present when math served as the outcome  $D(4, 293)=0.06, p<.01$ . Nevertheless, violations of the assumption of normality is generally not a concern in regression when the sample size is large (Field, 2015), and because of central limit theorem, large samples have repeatedly been shown to have normally distributed coefficient estimates, even if the error terms of the equation are not normally distributed (Berry, 1993, p. 82).

### **Domain General Contributions**

The goal of the first model used in this analysis was to identify age-related difference in the dynamic assessment of a domain general construct of working memory. In other words, the first model clustered together both verbal and visual measures of working memory when taken together to form a domain general construct. For Model 1a, a baseline model tested the contribution of initial working memory scores on reading and math achievement with ethnicity included as a covariate. In the subsequent models,

variables were added to determine their significance in predicting the achievement outcomes, and to evaluate the change in total  $R^2$  values. Also of interest was whether there is a unique contribution of each dynamic assessment measure when predicting reading and math achievement. In model 1b and 1c, the dynamic assessment measures were entered, with the gain score analyzed as part of model 1b, and the maintenance score in model 1c. Each of these steps also included age as a main effect as well as the interaction of age with the working memory measures (initial, gain, maintenance). Model 1d then tested the significance of all independent variables as well as age as a main effect and as an interaction effect with each of the three working memory scores.

### **Results of Domain General Models (Models 1a to 1d)**

Results of the domain general regression analysis can be found in Table 3. The results of model 1a indicated tests of working memory served as significant predictors of achievement in both reading  $F(4, 293)=20.89, p <.01, R^2 = .22$ , and math  $F(4, 293)=34.28, p <.01, R^2 = .32$ . The results in Table 3 show that age was a significant predictor of achievement after partialing out the shared variance of ethnicity and initial working memory scores. As expected, scores in reading and math improved with increases in age. An age x initial score interaction effect was noted to be significant only when predicting math  $F(1, 293)=4.80, p <.05$ , but not reading  $F(1, 293)=0.01, p=.94$ .

The addition of gain scores into model 2b was also considered significant for both reading  $F(4, 293)=29.37, p <.01, R^2 = .29$  and math  $F(4, 293)=42.31, p <.01, R^2 = .37$ . When considering age, it was noted that the main effect of age was only significant when math was selected as the outcome variable  $F(1, 293)=21.31, p <.01$ , and the interaction

between age and gain score was only significant  $F(1, 293)=7.52, p <.01$  when predicting math.

When maintenance scores were selected as the sole dynamic measure, the model was again found to be significant when predicting both reading  $F(4, 293)=26.75, p <.01, R^2 = .27$ , and math  $F(4, 293)=40.59, p <.01, R^2 = .37$ . As shown for the full model, Maintenance scores were a significant predictor of both reading  $F(1, 293)=15.93, p <.01, R^2 = .31$  and math  $F(1, 293)=22.90, p <.01, R^2 = .39$  after controlling for the shared variation predicted by both initial and gain scores.

To compare the base and final models, an F-ratio was calculated to evaluate the effect of change in predicted achievement between the full and base models. This was calculated using the formula in Tabachnick and Fidell (2013), where  $R^2_{wi}$  represents the associated  $R^2$  value of the full model,  $R^2_{wo}$  is the associated  $R^2$  value of the base model,  $m$  represents the number of new variables added to the full model from the base model, and  $df_{res}$  is the total residual degrees of freedom.

$$F = \frac{\left( R^2_{wi} - R^2_{wo} / m \right)}{\left( 1 - R^2_{wi} / df_{res} \right)}$$

When comparing models, significant differences in the  $R^2$  were noted between the full model and base models when predicting both reading  $F(4, 289)=8.33, p=.<.01$  and math  $F(4, 289)=8.29, p=.<.01$ , indicating the addition of dynamic assessment of working memory does contribute to the prediction of significant variance in achievement beyond the contribution of initial scores.

Taken together, the results of model 1 suggest that dynamic testing measures contribute unique variance in predictions of reading and math. As evidenced by the F-ratio calculated between the full and base models, the use of dynamic scores significantly contributed unique variance in predicting reading and math achievement scores independent of the initial, (non-dynamic) scores.

When the variables are partialled in the analysis, the results indicate achievement scores were moderated by the function of age on the initial/baseline scores, and the interaction between age and dynamic assessment of working memory predicted unique variation in math scores. To clarify these unique age-related interactions, a follow up ANCOVA was computed, and the results are discussed after the presentation of the full second model.

### **Modality Specific Contribution.**

While the previous analysis assessed the contribution of working memory as a domain general construct, the second series of analyses evaluated domain-specific variance related to modality (verbal vs. visual working memory) in predictions of reading and math performance. For model 2a, the baseline model tested the contribution of verbal and visual-spatial baseline working memory scores on both reading and math achievement. As noted in the subsequent models, variables were added to the model to identify their unique variance in predicting the achievement outcomes, and to evaluate the change in total  $R^2$  values.

In model 2b and 2c, the main effects of the dynamic assessment measures were entered, which included gain scores for verbal working memory and visual working

memory in model 2b, and maintenance scores for each modality in model 2c. Interaction effects pertaining to the modality (age x verbal and age x visual working memory) were also entered as well as age as a main effect. Model 2d then tested the significant contribution of all measures of working memory separated by modality (initial verbal and visual scores followed by gain verbal and visual, then maintenance verbal and visual), as well as age as a main effect and as an interaction effect between each of the initial and dynamic working memory scores.

### **Results of Domain Specific Models (Models 2a to 2d)**

Results of the domain specific regression models can be found in Table 4. The base models, which contained ethnicity, verbal, and visual initial working memory scores, as well as age as a main effect and as an interaction effect with both modalities, were significant when predicting reading  $F(6, 291)=14.88, p <.01, R^2 = .23$  and math  $F(6, 291)=23.58, p <.01, R^2 = .33$ . When accounting for age and the interaction effects with each domain, it was noted that the initial verbal domain was no longer a significant predictor of reading  $F(1, 291)=0.57, p =.45$ , or math  $F(1, 291)=0.36, p=.55$ . The initial visual domain, however, was a significant predictor of both reading  $F(1,291)=7.61, p <.01$ , and math,  $F(1,291)=13.91, p <.01$ . No significant interactions emerged between age and domain specific working memory scores in predicting either reading or math performance.

When substituting gain scores for the initial scores in regression model 2b, it was noted that the model was significant for predicting reading  $F(6,291)=20.18, p <.01, R^2 = .29$  and math  $F(6, 291)=28.36, p <.01, R^2 = .37$ . Regardless of achievement outcome, it

was also noted that visual gain scores were able to predict significant variation in the outcomes when accounting for the shared variation with verbal gain scores. As was the case in the baseline model (2a), no interaction effects were noted between age and domain specific gain scores.

Model 2c was also considered to be a significant predictor of reading  $F(6, 291)=19.01, p <.01, R^2 = .28$  and math  $F(6, 291)=27.31, p <.01, R^2 = .36$ . Here, the pattern of results mirrored those described in model 2b as ethnicity and both verbal and visual dynamic measures of working memory predicted unique variance in both reading and math. Similarly, age as a main effect was noted to be significant only when predicting math achievement  $F(1, 291)=23.26, p <.01$ . No significant interaction effects were noted between age and domain specific working memory maintenance scores when predicting math; however, the age x visual maintenance score was considered to be significant when predicting reading achievement  $F(1, 291)=-2.49, p <.05$ .

The full model (2d) was computed by entering all main effects and interaction effects for age, initial and dynamic assessment measures. The model was significant for both reading  $F(14, 283)=9.66, p <.01, R^2 = .32$ , and math  $F(14, 283)=13.97, p <.01, R^2 = .41$ . As was the case when analyzing the domain general model (Model 1d), the full model yielded a higher  $R^2$  value than the base model. When predicting reading, the full model generated an  $R^2$  value that was .09 greater than the initial/baseline model, and this increment in  $R^2$  was significant when compared to the baseline (Model 2a),  $F(8, 283)=4.58, p <.001$ . When predicting math scores, the full model  $\Delta R^2$  value was .08

higher than the baseline model, and this increase was also significant  $F(8, 283)=4.76$ ,  $p<.001$ .

As was the case in model 2c, age x visual maintenance score was a significant contributor to the prediction of variation in reading  $F(1, 283)=2.02$ ,  $p <.05$ , but was also significant in predicting math  $F(1, 283)=4.78$ ,  $p <.05$  scores. Taken together, the Full Model (2d) supports the notion that dynamic measures of working memory do predict additional variation in reading and math achievement beyond the contribution of initial scores. Based on the results of models one and two together, the null hypothesis to the first research question was not supported, as dynamic assessment (as defined by measures of gain and maintenance) of working memory contributed unique variance in predictions of reading and math achievement beyond the contribution of static or initial score (working memory performance without dynamic assessment) performance.

In regards to the interaction effects between age and dynamic measure, these were noted when the dynamic measures were separated by modality. Age by visual maintenance scores were considered significant predictors of achievement. The results suggest that the interaction effects related to age are the most prevalent when working memory is treated as a domain general construct, and when predicting math achievement. Thus, the results of this study indicate there is evidence sufficient to reject the null hypothesis associated with the second research question, as the effects of dynamic assessment (gain scores, maintenance scores) of working memory on math achievement outcomes are moderated by age.

When conceptualizing working memory into domain specific modalities, the results of this study suggests interaction effects are present when analyzing age and visual maintenance scores. This interaction predicted reading when maintenance was the only dynamic score in the model, and predicted math above and beyond the variation predicted by all other scores and their interaction with age. Thus, partial evidence for accepting the alternative hypothesis for research question three is present as the effects of dynamic assessment of working memory on reading and math achievement outcomes are moderated by the interaction between the dynamic assessment of visual working memory and age.

To summarize, after the variables are partialled in the analysis, and working memory is analyzed using domain specific tasks (e.g., visual and verbal), unique contributions were noted in the prediction of reading using gain and maintenance scores. Dynamic assessment of visual working memory (gain and maintenance) was noted to predict unique variation in reading and math after controlling for the effects of verbal working memory.

### **ANCOVA**

Analysis of covariance was then performed in cases where a unique interaction (partialled for the influence of other variables) between age and a dynamic test of working memory was found to be significant in the full first model. The one-way ANCOVA included five age groups as an independent variable and categorical measures of dynamic assessment (gain, maintenance). Participants were grouped by ages that identify their current general position in K-12 education (e.g., early elementary, late elementary, etc.).



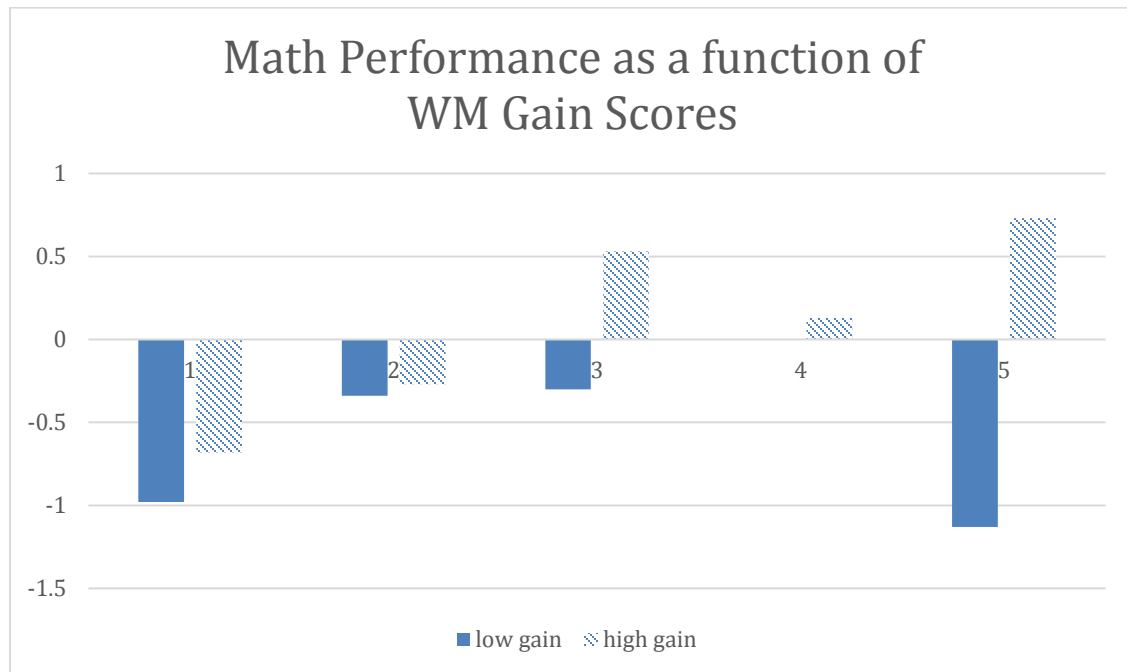
Dynamic assessment results were converted to classification measures for the ANCOVA. Results were classified as low (z-scores < 0) and high (z-scores > 0) performance. Covariates used in the model included initial scores and ethnicity. Post hoc comparisons were then used to identify which age brackets demonstrate the greatest benefit from dynamic assessment.

This analysis focused on a domain general conceptualization of working memory (Model 1) and its contribution to math, as the interaction effects in predicting reading were not significant. Because interaction effects were noted between age and both the gain and maintenance scores, two follow up analyses were run.

### **Gain Score Related Interactions**

For the first ANCOVA, independent variables included age group followed by dynamic gain scores classified into two groups (low and high gainers). To simplify the analysis, all reading and math scores were converted to z-scores based on the total sample. This was then followed by the interaction effect between age groups and dynamic gain score classifications. A median split was used to determine high and low gain scores. The covariates were initial working memory scores and ethnicity. From this analysis, the least square means were computed and are shown in Figure 1

Figure 1



As shown in Figure 1, there is an increase in z-scores for math as a function of age. However, the increase is isolated to participants who show a high gain in performance. As shown in Figure 1, the adjusted z-scores indicated that gainers in age groups 3 (10-11 years) and 5 (14-17 years) yielded the highest math scores. It is important to note that for age group 4, the least square mean for the low gain score could not be computed as the number of participants in this group was too low for statistical analysis. The Figure clearly shows, however, that gainers had higher math scores than nongainers for age group 5.

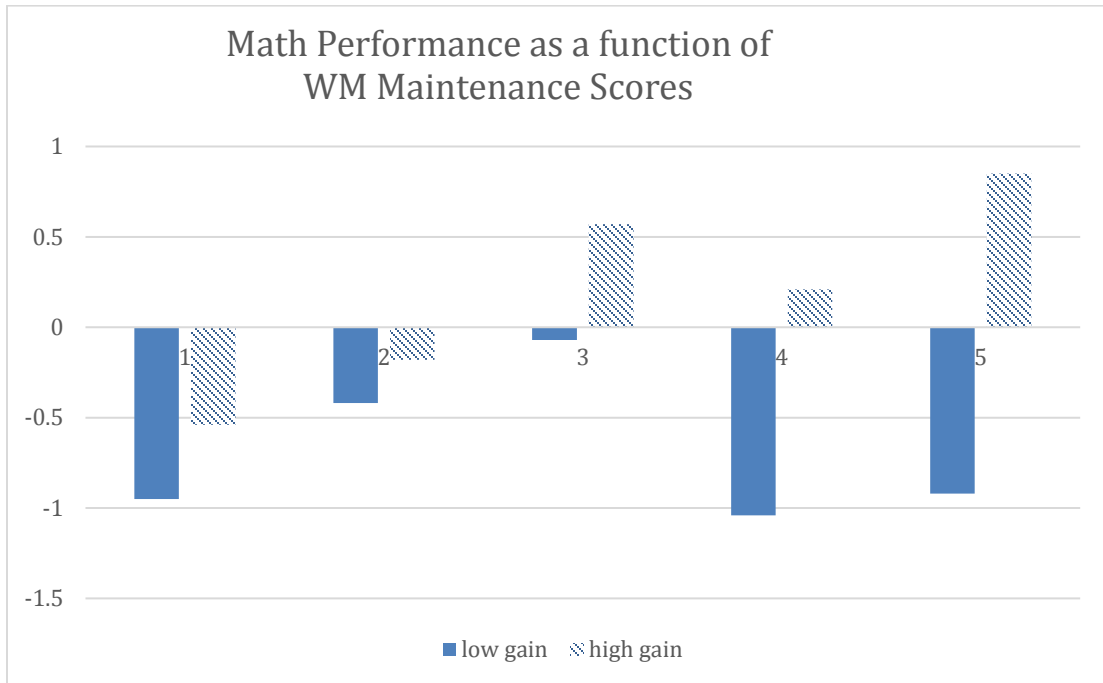
As expected from the previous regression analyses, the results of this ANCOVA model were considered to be significant  $F(16, 280)=12.83, p <.01, R^2 =.42$  when predicting math calculation scores. The main effect for high gainers (scores above or

below the median split) was also significant,  $F(1, 280)=12.85, p <.001$ , but as shown in Figure 1, a significant age x gain interaction emerged,  $F(4, 280)=4.92, p <.01$ . A post hoc Tukey test was used to gather additional information on the significant differences in least square means. The post-hoc comparisons of the least square means between gainers and nongainers as a function of group comparisons are displayed in Table 6. The analysis shows that high gainers from age groups 3, 4, and 5 demonstrated stronger math scores than the high gainers in age groups 1 and 2. When evaluating the low gainers, no significant differences were noted between the age groups.

### **Maintenance-Related Interactions**

A second follow-up ANCOVA was run to analyze the significant interaction effects associated with the maintenance scores. The covariates were initial scores and ethnicity. Age group and the dynamic assessment of working memory score (Maintenance), which was classified into two groups (low and high maintainers) were entered into the model. Results of the analysis found the model to be a significant predictor of math calculation skills  $F(16, 279)=13.34, p <.01, R^2 =.43$ .

Figure 2



Differences between and within the age groups are illustrated in Figure 2. As shown, participants who were labeled as high maintainers yielded higher math scores than participants labeled as low or nonmaintainers within each age group. Similar to the findings that used the gain score as the independent variable, the adjusted z-scores for the maintenance condition also indicated that maintainers in group 5 yielded the highest math scores. A Tukey test was used to identify the differences between maintainers within each age group. Results are displayed in Table 7. Similar to the previous ANCOVA, the results of this analysis indicated the use of dynamic assessment of working memory (maintenance score) with 10 to 11-year-old students as well as 14 to 17-year-olds was more effective in identifying variation in math achievement scores when compared to

both five to seven and eight to nine-year-old participants, while high gainers from age group 4 outperformed the high gainers from age group 1.

Taken together, this follow up analysis indicates that the largest impact in math performance between maintainers and non maintainers occurred for age groups 3 through 5 (10-17 years of age).

## Chapter 5 - Discussion

The purpose of this study was to explore age-related differences in the dynamic assessment of working memory when predicting reading and math outcomes. Exploration of this issue involved defining working memory as both a domain general and a domain specific construct, thus models were run to analyze potential interaction effects using both conceptualizations of working memory. Participants in the study were taken from a secondary data set, which included 602 individuals, although this sample was reduced to 298 participants due to missing data on the outcome measures. All individuals were evaluated between 1990 and 2012. Examiners were graduate students from primarily five universities, and all examiners were trained to administer standardized assessments, with specific instruction given in administering and scoring the dependent and independent measures.

This study sought to answer three research questions. First, the effectiveness of dynamic assessment of working memory in predicting variation in reading and math achievement over and above that which is predicted by initial scores was explored using two conceptualizations of working memory: a domain general construct, which considered measures of verbal and visual working memory together, and a domain specific construct which separate these modalities for individual analysis. The second research question asked if interaction effects between age and dynamic assessment of working memory (domain general) predicted variation in reading and math, while the third question reframed the previous question using a domain specific model of working

memory (verbal versus visual measures). Results pertaining to each of these questions is summarized below.

**1. Does the dynamic assessment (as defined by measures of gain, maintenance) of working memory contribute unique variance in predictions of reading and math achievement beyond the contribution of static or initial score (working memory performance without dynamic assessment) performance?**

Results of model one provide evidence suggesting dynamic assessment of working memory is able to predict additional variation in reading and math achievement beyond that predicted by static assessment. As evidenced by the calculated F ratio between the final and base models, dynamic scores predict more variation in both reading ( $\Delta R^2 = .09$ ) and math ( $\Delta R^2 = .07$ ) achievement scores than when initial, non-dynamic scores are used.

When visual and verbal modalities were separated, the pattern of results continued to support the capacity for dynamic measures of working memory to predict unique variation in achievement. Similar increases in  $R^2$  values were noted. These findings are considered to be commensurate with the results of previous research (Swanson, 1992; 1999; 2010; 2011). Taken together, there is sufficient evidence to reject the null hypothesis as dynamic assessment of working memory does appear to predict unique variance in reading and math scores after controlling for baseline working memory ability.

**2. Are the effects of dynamic assessment (gain scores, maintenance scores) of working memory on reading and math achievement outcomes moderated by age?**

Based on the results of model one, the effects of dynamic assessment of working memory on math do appear to be moderated in part by age. This effect was noted when reviewing the interaction between the gain score and age, and the interaction between age and maintenance while controlling for the effects of the initial and gain score respectively, as well as age as a main effect. Interactions between age and dynamic assessment was limited to the prediction of math as no significant interaction was noted when reading was used as the outcome. Thus, the results of this study indicate there is evidence sufficient to reject the null hypothesis associated with the second research question, but acceptance of the alternative hypothesis is limited to the use of math as the outcome variable.

**3. Are the effects of dynamic assessment of working memory on reading and math achievement outcomes moderated by the interaction between the modality type of dynamic assessment (visual, verbal) measure and age?**

When considering modality specific measures, only age by visual maintenance scores were considered significant predictors of achievement. Interactions between visual maintenance scores and age were significant predictors of reading achievement when the maintenance score was the only dynamic working memory score used (model 2c), and in predicting reading and math when controlled for the shared variation predicted by the initial, gain, and maintenance scores, as well as age and interactions between age and



initial and gain scores. With clear interactions noted based on different modalities, the null hypothesis for question three is rejected; however, because interaction effects were generally limited to measures involving visual maintenance, the alternative hypothesis is conditionally accepted.

### **Implications**

*Improvements in the prediction of achievement:* Various arguments could be made which explain why dynamic assessment of working memory is able to predict more variation in achievement scores when compared to baseline ability. This author believes there are two primary reasons why the dynamic measures in the S-CPT/aTWM serve as stronger predictors of achievement. First, prompting during the dynamic procedure reduces poor performance in working memory scores related to encoding. The ability to hold and manipulate information in working memory can only be accomplished if an individual has first encoded that information into memory. It has been documented in the literature that children who struggle with encoding of information have a tendency to improve their performance when the information is repeated (Wilson, 2009). Because the prompting used in this assessment procedure involves selective reminding of information in the initial, followed by the final and medial position, and ultimately a complete repetition of the entire item, it may be that during the dynamic assessment process, a purer measure of working memory capacity is acquired because the examiner is able to improve the initial encoding. Such a hypothesis is in line with Swanson's (2007) explanation of the gain score relating to processing efficiency and the maintenance score associated with working memory capacity.

A second explanation for why dynamic assessment of working memory is a better predictor of achievement lies in the prompting for strategy selection. During the teaching step of the assessment, the examiner provides the participant with an opportunity to choose from four strategies for executing the task. One may argue that selection of a strategy for completing a working memory task is usually controlled by the central executive (Barkley, 2012; McCloskey & Perkins, 2013). Because the examiner provides such options to the examinee, it may be that he or she is cuing the examinee's central executive to work more efficiently.

In addition, significant variation in executive skills has been noted between individuals. Borkowski and Burke (2005) indicated that the degree one utilizes executive functions to solve a problem is dependent upon his or her intelligence, as more intelligent individuals are better equipped to draw on lower order skills to solve a task, while individuals with weaker intelligence must draw on the executive system to a greater degree. Providing cues for strategy selection and improving encoding may adjust performance on the independent variables in this study so that variation in the executive skills between participants is reduced; therefore, the dynamic score serves a purer measure of working memory capacity.

*Interactions between age and dynamic assessment of working memory:* As indicated by the results of this study, the added benefit of dynamic assessment varies between age groups; however, this interaction effect was noted to be a significant predictor of math, but not reading. In addition, the interaction appears to be occurring between age and visual spatial measures of working memory. These effects raise several

questions. First, why are age related differences noted in the added benefit of dynamic assessment of visual spatial working memory? The answer to this may lie in the developmental trajectory of the central executive (mainly attentional control), and age-related differences in visual spatial memory span.

First, one can argue that visual spatial memory capacity is generally poor in younger children but improves with age (Cowan, Morey, AuBuchon, Zwillling, Gilchrist, 2010; Shimi, Nobre, Astle, & Scerif, 2014). Dramatic increases in memory capacity have been noted between the ages of three and 10 years (Gathercole, 1999). When memory span is considered to be poor, the added benefit of the dynamic process, which serves as a selective reminder or cue to better encode information into short term memory, may be less effective because the nature of the dynamic tasks improve encoding, but are less geared towards equipping the examinee with memory aides to improve storage. In other words, if storage is poor, the impact of additional cues to encoding information into storage is limited. Older individuals, however, show improvements in their overall memory capacity; therefore, the prompts provided to attain the gain scores may transferred to improvements in holding the information within the mind's eye.

Second, previous research has suggested that the ability to orient attention towards a visual spatial task improves the encoding and storage in visual short-term memory (Baddeley & Hitch, 1974; Engle, Tuholski, Laughlin, & Conway, 1999; Griffin & Nobre, 2003; Makovski & Jiang, 2007). Cueing an individual to remember visual spatial material appears beneficial regardless of one's age, even in younger children who typically demonstrate poor overall memory (Astle, Nobre, & Scerif, 2012). For this

reason, individual differences in attentional control appear to be an important variable which predicts variation in visual memory capacity (Astle, et al., 2012). Given that the role of the central executive in the working memory system has been identified as an important component when predicting achievement (Bull, Johnson, & Roy; 1999; Geary, Hoard, Byrd-Craven, Nugent, & Numtee, 2007; Swanson, 1993), one may expect individual differences in this area to play an important part in predicting variation in math calculation skills.

Research has demonstrated that when external cues are presented to assist an individual in orienting his or her attention to a visual spatial memory task after the initial presentation of the information to be remembered, the representation of that information is improved during the maintenance period (Raye, Johnson, Mitchell, Greene, & Johnson, 2007; Shimi, Nobre, Astle, & Scerif, 2014). Shimi and colleagues (2014) found that while the memory for visual spatial information in seven-year-old children did show benefit from cues to allocate their attentional resources towards the material, 11-year-old children and adults were shown to have a greater ability to do so when the cues were presented after the initial presentation of information. In addition, they found that voluntary rather than involuntarily-controlled attention was responsible for the benefits and the age-related differences. Based on their findings, Shimi et al. suggested that older individuals are in a better position to allocate visual attentional resources in order to update information held in the mind's eye, while inhibiting distracting information, and the allocation of attentional resources appears to be an important component when evaluating the differences noted across age groups. Furthermore, the need to inhibit

distracting information may also explain age related differences in the dynamic assessment of working memory as this skill does not appear to reach maturity until age 10 (Maricle, Johnson, & Avirett, 2010).

A second question one may ponder based on the findings of this study is why does the interaction between age and the dynamic assessment of working memory predict math achievement, but not reading? As noted above, attentional control is unlikely to be the sole explanation for the interaction effects noted in the current study, as this executive skill has often been conceptualized as a domain-general resource (Baddeley & Hitch, 2000). If developmental differences in attentional control was the single explanation for the interaction effect with age when predicting achievement, then one would likely expect these differences to also predict variation in reading. Because this is not the case, the relationship between developmental differences in attentional control and increases in visual spatial memory span seem a more suitable explanation, as visual spatial working memory has a stronger relationship with math than it does with reading. Such an explanation would align with the findings of Cowan et al. (2010) who reasoned that both the ability to allocate attentional resources and limited capacity for visual storage were important factors in explaining developmental differences in visual spatial short-term memory.

A third question one may consider is why do children in the older age groups (10 to 11, 12 to 13, and 14 to 17-years-old) demonstrate added benefit from dynamic assessment of working memory when predicting math over the two youngest age groups in this study (five to seven and eight to nine-year-old students)? Perhaps this question can

be answered by evaluating the relationship between working memory and math achievement across development. The importance of each of the slave systems under the central executive in Baddeley's (2000) model of working memory and their relationship to predicting variation in math achievement certainly varies depending on the complexity of the math (Bull et al., 2008; Geary, 2013; Geary et al., 2007). The phonological loop, for example, is important in the retrieval of basic math facts from long-term memory, while the visuospatial sketchpad is more involved in a wider range of math achievement (Fuchs et al., 2006; Geary, 1993; 2013). Bull and colleagues (2008) identified visuospatial short-term memory as an important predictor of math achievement in young preschool children, and indicated visual spatial working memory tasks became a stronger predictor around the ages of seven to eight years.

If the cognitive demands shift from a stronger reliance on the phonological loop to a higher need to use one's visuospatial sketchpad, and an age-related increase in the need for attentional control is also present, then the dynamic visual working memory measures discussed in this paper would likely serve as stronger predictors of math in the older participants because they cue the participant to use these skills. Younger students rely more on the phonological loop to solve math problems; thus, prompting more efficient use of the visuospatial sketchpad is less important/effective.

In summary, the author believes dynamic assessment may be a better predictor of achievement because it improves encoding into memory, and because it includes prompting to use a strategy. Prompts to select a strategy may cue the examinee to more efficiently use his or her central executive, which can allow for a more pure measure of

working memory capacity. Age-related differences in the added benefit of dynamic assessment of working memory when predicting math were noted, and may be related to the interactions between the progressive demands that math curriculum places on working memory systems, the growing importance of the visuospatial sketchpad in completion of math calculations as one ages, and the development of attentional control. This explanation is presented as a possibility that appears to align with current understanding of math achievement, working memory, and executive functioning; however, a direct analysis of this explanation was not included in the current study. Additional research would be needed to evaluate the validity of this reasoning.

### **Limitations of the study**

Limitations in this study are present and should be considered prior to generalizing findings. First, this study focuses exclusively on the potential added benefit in terms of predictive power. As noted in the literature review, one important application of dynamic assessment is to assist in the development of interventions for a child, and this application is not considered as part of this study. The differences in predictive capacity discussed in this study do not translate to age-related differences in predicting potential intervention strategies. Additional research would be needed to explore this area.

A second limitation of the current study lies in the methods of data collection. Although it was noted that each examiner was trained to administer standardized assessments in general, and given specific instruction to administer the dependent and independent variables in this study, there were no methods established to determine inter-

rater reliability. Using more than one rater at the time of test administration and response recording would have been a major logistic challenge. Although the use of multiple raters when administering and scoring a single assessment is generally absent in the standardization procedures used by many of the major test publishers, it is nevertheless a source of unexplained error and should be considered when interpreting results.

For the purpose of this study, outliers were defined as any score which fell greater than 3 standard deviations from the mean. In relation to the outcome variables, only one outlier was found on the mathematics assessment, and was retained. When evaluating the distribution of scores on the independent measures, there were several outliers present, and many of these outliers were noted in children between the ages of 10 and 11 years.

Acknowledgement of these outliers is necessary; however, they are not considered to be a significant contributor to the findings discussed in this study. Upon identification of these outliers, the data were first analyzed with all participants. After obtaining results, the researcher elected to complete a list wise deletion of all participants who had z-scores falling more than three standard deviations from the mean. The data was then reanalyzed and compared to the initial results. Deletion of the outliers did not have a significant impact on the overall findings of the study as dynamic assessment still improved prediction of reading and math, and age by dynamic assessment interactions were still present in the areas discussed in this paper. Of course, differences in the beta values were present, and this complicates the translation of these findings into a prediction model, but otherwise, the main findings within this study were consistent. Based on this comparison, the author elected to retain all data. In large part, this decision



was based on the conceptualization of dynamic assessment. The author believes variation in the zone of proximal development should be embraced, and, while small to non-existent changes in performance or even major improvements in performance based on scaffolding may be statistical anomalies, they still lie at the heart of dynamic assessment and have possible implications for the individual student.

It was noted throughout the result section that high gainers and high maintainers performed better on measures of math calculation than their younger peers, but this was generally not the case for students who were low gainers/maintainers. While it could be reasoned that an inability to show improvement in the dynamic assessment process tends to predict low math scores regardless of age, it must be noted that the sample size within the cells for older, low gainers and maintainers was small. For this reason, an alternative explanation may be that the small sample size within these cells was unable to generate the necessary power to identify the presence of significant differences.

Potential threats to instrumentation were also noted. This is a common concern with studies involving dynamic assessment, as repeated exposure to a task can change an individual's performance. In the case of this study, which includes measures of working memory and probes that request the participant to select a strategy to help him or her remember, this concern is forefront. Many definitions of executive functioning include organizing performance on *novel* tasks (McCloskey & Perkins, 2013). In addition, fMRI research indicates that with repeated exposure to a task along with opportunities to practice, the executive demands tend to decrease, allowing for posterior parts of the brain to process the problem; even when participants are exposed to a similar task constructed

with novel items, the allocation of cognitive resources away from the frontal lobes tend to remain (Posner & Raichle, 1994). This suggests that the construct of the skill does indeed change. Yet such a concern for this study should be embraced rather than viewed with caution. After all, creating such a change within the testing environment, although problematic for quantitative analysis, is one of the very reasons why we do dynamic assessment.

Use of a specific measure of processing rather than selecting a broad range of tests may narrow the external validity of this study. Inclusion of additional independent measures, however, was not considered practical as the S-CPT and aTWM are currently the only published dynamic tests of working memory that have adequate psychometric properties. Similarly, use of multiple outcome measures would have enriched the study's findings, although this was not considered because the current evaluation utilized preexisting data, and the administration of additional outcome measures to a sample size this large was not possible. It should also be noted that the current study was limited to the dynamic assessment of working memory, and age-related interaction effects should not be generalized to other forms of interactive assessment.

In conclusion, dynamic assessment of working memory is a useful procedure which provides an additional source to explain variation in achievement. In addition, age-related differences in this procedure do appear present and have potentially important implications for both practitioners and researchers to consider. This study demonstrated that dynamic assessment of working memory predicts additional variation in reading and math above that which is predicted by the baseline scores. These predictions are in part

moderated by the interaction between age and dynamic measure. When modeling a domain general construct of working memory, the interaction effect is noted when predicting math, but not reading, and it is noted with both the gain and maintenance scores. When working memory is divided into domain specific modalities, the interaction appears to only be present when using visual working memory maintenance scores, and was noted to be significant when predicting both reading and math. The limitations of this study should be considered prior to generalizing results.

## References

- Alloway, T. P., & Alloway, R. G. (2013). Working memory across the lifespan: A cross-sectional approach. *Journal of Cognitive Psychology, 25*, 84–93.
- Aravena, S., Tijms, J., Snellings, P., & van der Molen, M. W. (2016). Predicting responsiveness to intervention in dyslexia using dynamic assessment. *Learning and Individual Differences, 49*, 209-215.
- Astle, D. E., Nobre, A. C., & Scerif, G. (2012). Attentional control constrains visual short-term memory: Insights from developmental and individual differences. *Quarterly Journal of Experimental Psychology, 65*, 277–294.
- Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences, 4*, 417– 423.
- Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology, 63*, 1–29.
- Baddeley, A., & Hitch, G. (1974). Working memory. *Psychology of Learning and Motivation, 8*, 47-89.
- Baddeley, A. D., & Hitch, G. J. (2000). Development of working memory: Should the Pascual-Leone and the Baddeley and Hitch models be merged? *Journal of Experimental Child Psychology, 77*, 128–137.
- Baddeley, A. D., & Logie, R. H. (1999). The multiple-component model. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 28–61). Cambridge, United Kingdom: Cambridge University Press.
- Barkley, R. A. (2012). *Executive Functions: What they are, how they work, and why they evolved*. The Guilford Press: New York.
- Barrouillet, P., & Camos, V. (2001). Developmental increase in working memory span: Resource sharing or temporal decay? *Journal of Memory and Language, 45*, 1–20.
- Bayliss, D. M., Jarrold, C., Baddeley, A. D., Gunn, D. M., & Leigh, E. (2005). Mapping the developmental constraints on working memory span. *Developmental Psychology, 41*, 579 –597.
- Berry, W. D. (1993). *Understanding regression assumptions*. London: Sage Publications.

- Boers, E., Janssen, M. J., Minnaert, A. E., & Ruijsenaars, W. A. (2013). The application of dynamic assessment in people communicating at a prelinguistic level: A descriptive review of the literature. *International Journal of Disability, Development and Education*, *60*, 119-145.
- Borella, E., & de Ribaupierre, A. (2014). The role of working memory, inhibition, and processing speed in text comprehension in children. *Learning and Individual Differences*, *34*, 86–92.
- Borkowski, J. G., & Burke, J. E. (2005). Theories, models, and measurement of executive functioning: An information processing perspective. In G. R. Lyon & N. A. Krasnegor (Eds.), *Attention, memory, and executive function* (4th ed.). Baltimore, MD: Paul H, Brookes Publishing.
- Bridges, M. S., & Catts, H. W. (2011). The use of a dynamic screener of phonological awareness to predict risk for reading disabilities in kindergarten children. *Journal of Learning Disabilities*, *44*, 330–338.
- Budoff, M. (1987). *Measures for assessing learning potential*. In C. S. Lidz (Ed.), *Dynamic assessment* (pp. 82-115). New York: Guilford.
- Bull, R., Espy, K. A., & Wiebe, S. A. (2008). Short-term memory, working memory, and executive functions in preschoolers: Longitudinal predictors of mathematical achievement at age 7 years. *Developmental Neuropsychology*, *33*, 205-228.
- Bull, R., Johnson, R. S., & Roy, J. A. (1999). Exploring the roles of the visual-spatial sketchpad and central executive in children's arithmetical skills: Views from cognitive and developmental neuropsychology. *Developmental Neuropsychology*, *15*, 421-442.
- Caffrey, E., Fuchs, D., & Fuchs, L. S. (2008). The predictive validity of dynamic assessment: A review. *Journal of Special Education*, *41*, 254-270.
- Campione, J. C., & Brown, A. L. (1987). *Linking dynamic assessment with school achievement*. In C. S. Lidz (Ed.), *Dynamic assessment* (pp. 173-195). New York: Guilford.
- Carlson, J. S., & Wiedl, K. H. (1979). Towards a differential testing approach: Testing-the-limits employing the Raven Matricies. *Intelligence*, *3*, 323-344.
- Carlson, J. S., & Wiedl, K. H. (1992). *The dynamic assessment of intelligence*. In H. C. Haywood and D. Tzuriel (Eds.), *Interactive Assessment*. Springer-Verlag: New York.

- Caviola, S., Mammarella, I. C., Lucangeli, D., & Cornoldi, C. (2014). Working memory and domain-specific precursors predicting success in learning written subtraction problems. *Learning and Individual Differences, 36*, 92–100.
- Cho, E., Compton, D. L., Fuchs, D., Fuchs, L. S., & Bouton, B. (2014). Examining the predictive validity of dynamic assessment of decoding to forecast response to tier 2 intervention. *Journal of Learning Disabilities, 47*, 409-423.
- Conlin, J. A., Gathercole, S. E., & Adams, J. W. (2005). Children’s working memory: Investigating performance limitations in complex span. *Journal of Experimental Child Psychology, 90*, 303–317.
- Cormier, P., Carlson, J. S., & Das, J. P. (1990). Planning ability and cognitive performance: The compensatory effects of a dynamic assessment approach. *Learning and Individual Differences, 2*, 437-449.
- Coventry, W. L., Byrne, B., Olson, R. K., Corley, R., & Samuelsson, S. (2011). Dynamic and static assessment of phonological awareness in preschool: A behavior-genetic study. *Journal of Learning Disabilities, 44*, 322-329.
- Cowan, N., Morey, C. C., AuBuchon, A. M., Zwillling, C. E., & Gilchrist, A. L. (2010). Seven-year-olds allocate attention like adults unless working memory is overloaded. *Developmental Science, 13*, 120–133.
- Cronbach, L. J., & Furby, L. (1970). How we should measure “change” – or should we? *Psychological Bulletin, 74*, 68-80.
- Daneman, M., & Hannon, B. (2001). Using working memory theory to investigate the construct validity of multiple-choice reading comprehension tests such as the SAT. *Journal of Experimental Psychology: General, 130*, 208–223.
- Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review, 3*, 422–433.
- Denckla, M. B. (2005). A theory and model of executive function: A neuropsychological perspective. In G. R. Lyon & N. A. Krasnegor (Eds.), *Attention, memory, and executive function* (4th ed.). Baltimore, MD: Paul H, Brookes Publishing.
- Donaldson, A. L., & Olswang, L. B. (2007). Investigating requests for information in children with autism spectrum disorders: Static versus dynamic assessment. *Advances in Speech Language Pathology, 9*, 297–311.

- Embretson, S. E. (1987). Toward development of a psychometric approach. In C. L. Lidz (Ed.), *Dynamic Assessment: An interactional approach to evaluating learning potential*. Guilford Press: New York.
- Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General*, *128*, 309–331.
- Engle, R. W., Cantor, J., & Carullo, J. J. (1992). Individual differences in working memory and comprehension. A test of four hypotheses. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 972-992.
- Feuerstein, R., Rand, Y., & Hoffman, M. B. (1979). *The dynamic assessment of retarded performers: The learning Potential Assessment Device, theory, instruments, and techniques*. Baltimore, MD: University Park Press.
- Field, A. (2016). *Discovering statistics using IBM SPSS Statistics* (4th ed.). Los Angeles: Sage Publications.
- Field, A., & Miles, J. (2010). *Discovering statistics using SAS*. Los Angeles: Sage Publications.
- Flanagan, D. P., Alfonso, V. C., & Mascolo, J. T. (2006). A CHC-based operational definition of SLD: Integrating multiple data sources and multiple data gathering methods. In D. P. Flanagan and V. C. Alfonso (Eds.), *Essentials of specific learning disability identification*. Hoboken, NJ: Wiley Press.
- Fuchs, L. S., Compton, D. L., Fuchs, D., Hollenbeck, N., Craddock, C. F., & Hamlett, C. L. (2008). Dynamic assessment of algebraic learning in predicating third graders' development of mathematical problem solving. *Journal of Educational Psychology*, *100*, 829–850.
- Fuchs, L. S., Compton, D. L., Fuchs, D., Hollenbeck, Hamlett, C. L., & Seethaler, P. M. (2011). Two-stage screening for math-problem solving difficulty using dynamic assessment of algebraic learning. *Journal of Learning Disabilities*, *44*, 372-380.
- Fuchs, L. S., Fuchs, D., Compton, D. L., Powell, S. R., Seethaler, P. M., Capizzi, A. M., et al. (2006). The cognitive correlates of third grade skill in arithmetic, algorithmic computation, and arithmetic word problems. *Journal of Educational Psychology*, *98*, 29-43.
- Gathercole, S. E. (1998). The development of memory. *Journal of Child Psychology and Psychiatry*, *39*, 3–27.

- Gathercole, S. (1999). Cognitive approaches to the development of short-term memory. *Trends in Cognitive Sciences*, 3, 410–419.
- Gathercole, S. E., Lamont, E., & Alloway, T. P. (2006). Working memory in the classroom. In S. Pickering (Ed.), *Working memory in the classroom* (pp. 220 – 238) Oxford, England: Academic Press.
- Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Developmental Psychology*, 40, 177–190.
- Geary, D. C. (1993). Mathematical disabilities: Cognitive, neuropsychological, and genetic components. *Psychological Bulletin*, 114, 345-362.
- Geary, D. C. (2013). Learning disabilities in mathematics: Recent advances. In H. L. Swanson, K. R. Harris, & S. Graham (Eds.), *Handbook of Learning Disabilities* (2nd ed.). (pp. 239-255). New York: Guilford Press.
- Geary, D. C., Hoard, M. K., Byrd-Craven, J., Nugent, L., & Numtee, C. (2007). Cognitive mechanisms underlying achievement deficits in children with mathematical learning disability. *Child Development*, 78, 1343-1359.
- Gillam, S. L., Fargo, J., Foley, B. & Olszewski, A. (2011). A nonverbal phoneme deletion task administered in a dynamic assessment format. *Journal of Communication Disorders*, 44, 236-245.
- Glaspey, A. M., & MacLeod, A. A. N. (2010). A multi-dimensional approach to gradient change in phonological acquisition: A case study of disordered speech development. *Clinical Linguistics & Phonetics*, 24, 283–299.
- Glaspey, A. M., & Stoel-Gammon, C. (2007). A dynamic approach to phonological assessment. *Advances in Speech-Language Pathology*, 9, 286–296.
- Griffin, I. C., & Nobre, A. C. (2003). Orienting attention to locations in internal representations. *Journal of Cognitive Neuroscience*, 15, 1176–1194.
- Grigorenko, E. L. (2009). Dynamic assessment and response to intervention. Two sides of one coin. *Journal of Learning Disabilities*, 42, 111-132.
- Grigorenko, E. L., & Sternberg, R. J. (1998). Dynamic testing. *Psychological Bulletin*, 124, 75-111.



- Guthke, J., & Wingenfeld, S. (1992). The learning test concept: origins, state of the art, and trends. In H. C. Haywood and D. Tzuriel (Eds.), *Interactive assessment* (pp. 64-93). New York: Springer-Verlag.
- Gutierrez-Clellen, V. F., & Peña, E. (2001). Dynamic assessment of diverse children: A tutorial. *Language, Speech, and Hearing Services in Schools, 32*, 212–224.
- Hall, D., Jarrold, C., Towse, J. N., & Zarandi, A. L. (2015). The developmental influence of primary memory capacity on working memory and academic achievement. *Developmental Psychology, 51*, 1131–1147.
- Hamilton, C. J., Coates, R. O., & Heffernan, T. (2003). What develops in visuo-spatial working memory development? *European Journal of Cognitive Psychology, 15*, 43–69.
- Haywood, H. C., & Tzuriel, D. (2002). Applications and challenges in dynamic assessment. *Peabody Journal of Education, 77*, 40–63.
- Holmes, J., Adams, J. W., & Hamilton, C. J. (2008). The relationship between visuospatial sketchpad capacity and children's mathematical skills. *European Journal of Cognitive Psychology, 20*, 272–289.
- Jastak, S., & Wilkinson, G. S. (1984). *The Wide Range Achievement Test-Revised*. Wide Range, Inc.
- Jitendra, A. K., & Kameenui, E. J. (1993). Dynamic testing as a compensatory testing approach: A description and analysis. *Remedial and Special Education, 14*, 6-18.
- Kantor, P. T., Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (2011). Comparing two forms of dynamic assessment and traditional assessment of preschool phonological awareness. *Journal of Learning Disabilities, 44*, 313-321.
- Kar, B. C., Dash Utkal, U. N., Das, J. P., & Carlson, J. (1993). Two experiments on the dynamic assessment of planning. *Learning and Individual Differences, 5*, 13-29.
- Kirkwood, M. W., Weiler, M. D., Bernstein, J. H., Forbes, P. W., & Waber, D. P. (2001). Sources of poor performance on the Rey-Osterrieth Complex Figure Test among children with learning difficulties: A dynamic assessment approach. *The Clinical Neuropsychologist, 15*, 345-356.
- Kong, J. E., & Orosco, M. J. (2016). Word-problem-solving strategy for minority students at risk for math difficulties. *Learning Disability Quarterly, 39*, 171-181.

- Kozulin, A., & Garb, E. (2002). Dynamic assessment of EFL text comprehension. *School Psychology International, 23*, 112-127.
- Luciana, M., Conklin, H. M., Hooper, C. J., & Yarger, R. S. (2005). The development of nonverbal working memory and executive control processes in adolescents. *Child Development, 76*, 697–712.
- Makovski, T., & Jiang, Y. V. (2007). Distributing versus focusing attention in visual short-term memory. *Psychonomic Bulletin & Review, 14*, 1072–1078.
- Maricle, D. E., Johnson, W., & Avirett, E. (2010). Assessing and intervening in children with executive function disorders. In D. C. Miller (Ed.). *Best practices in school neuropsychology: Guidelines for effective practice, assessment, and evidence-based intervention*. Hoboken, NJ: John Wiley & Sons.
- McAuley, T., & White, D. A. (2011). A latent variables examination of processing speed, response inhibition, and working memory during typical development. *Journal of Experimental Child Psychology, 108*, 453–468.
- McCabe, J., & Hartman, M. (2003). Examining the locus of age effects on complex span tasks. *Psychology and Aging, 18*, 562–572.
- McCloskey, G., & Perkins, L. A. (2013). *Essentials of executive functions assessment*. Hoboken, NJ: Wiley & Sons.
- McGrew, K., & Wendling, B. (2010). Cattell-Horn-Carroll cognitive achievement relations: What we have learned from the past 20 years of research. *Psychology in the School, 47*, 651-675.
- McLaughlin, K., & Cascella, P. W. (2008). Eliciting a distal gesture via dynamic assessment among students with moderate to severe intellectual disability. *Communication Disorders Quarterly, 29*, 75-81.
- Murphy, R., & Maree, D. J. F. (2006). A review of South African research in the field of dynamic assessment. *South African Journal of Psychology, 36*, 168-191.
- Nigam, R. (2001). Dynamic assessment of graphic symbol combinations by children with autism. *Focus on Autism and other Developmental Disabilities, 16*, 190–197.
- O'Connor, R. E., & Jenkins, J. R. (1999). Prediction of reading disabilities in kindergarten and first grade. *Scientific Studies of Reading, 3*, 159-197.

- Orosco, M. J., Swanson, H. L., O'Connor, R., & Lussier, C. (2013). The effects of dynamic strategic math on English language learners' word problem solving. *The Journal of Special Education, 47*, 96–107.
- Peterson, D. B., Allen, M. M., and Spencer, T. D. (2014). Predicting reading difficulty in first grade using dynamic assessment of decoding in early kindergarten: A large-scale longitudinal study. *Journal of Learning Disabilities, 49*, 200-215.
- Pishghadam, R., Barabadi, E., & Kamrood, A. (2011). The differing effect of computerized dynamic assessment of L2 reading comprehension on high and low achievers. *Journal of Language Teaching and Research, 2*, 1353-1358.
- Posner, M. I., & Raichle, M. E. (1994). *Images of Mind*. New York: Scientific American Library
- Raye, C. L., Johnson, M. K., Mitchell, K. J., Greene, E. J., & Johnson, M. R. (2007). Refreshing: A minimal executive function. *Cortex, 43*, 135–145.
- Saldana, D. (2004a). Dynamic Master Mind: Interactive use of a game for testing metacognition. *School Psychology International, 25*, 422-438.
- Saldana, D. (2004b). Interactive assessment of metacognition: Exploratory study of a procedure for persons with severe mental retardation. *European Journal of Psychology of Education, 19*, 349-364.
- Schroeder, P. J. (2014). The effects of age on processing and storage in working memory span tasks and reading comprehension. *Experimental Aging Research, 40*, 308–331.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Belmont, CA: Wadsworth Cengage Learning.
- Shimi, A., Nobre, A. C., Astle, D., & Scerif, G. (2014). Orienting attention within visual short-term memory: Development and mechanisms. *Child Development, 85*, 578-592.
- Snell, M. E. (2002). Using dynamic assessment with learners who communicate nonsymbolically. *Augmentative and Alternative Communication, 18*, 163–176.
- Spector, J. E. (1992). Predicting progress in beginning reading: Dynamic assessment of phonemic awareness. *Journal of Educational Psychology, 84*, 353-363.
- Swanson, H. L. (1992). Generality and modifiability of working memory among skilled and less skilled readers. *Journal of Educational Psychology, 84*, 473-488.

- Swanson, H. L. (1993). Working memory in learning disability subgroups. *Journal of Experimental Child Psychology*, 56, 87-114.
- Swanson, H. L. (1995). *Swanson-Cognitive Processing Test (S-CPT)*. Austin, TX: Proed.
- Swanson, H. L. (1996). Classification and dynamic assessment of children with learning disabilities. *Focus On Exceptional Children*, 28, 1-20.
- Swanson, H. L. (1999). What develops in working memory? A lifespan perspective. *Developmental Psychology*, 35, 986-1000.
- Swanson, H. L. (2003). Age-related differences in learning disabled and skilled readers' working memory. *Journal of Experimental Child Psychology*, 85, 1-31.
- Swanson, H. L. (2008). Working memory and intelligence in children: What develops? *Journal of Educational Psychology*, 100, 581-602.
- Swanson, H. L. (2010). Does the dynamic testing of working memory predict growth in nonword fluency and vocabulary in children with reading disabilities? *Journal of Cognitive Education and Psychology*, 9, 139-165.
- Swanson, H. L. (2011). Dynamic testing, working memory, and reading comprehension growth in children with reading disabilities. *Journal of Learning Disabilities*, 44, 358-371.
- Swanson, H. L. (2013). *Abbreviated test of working memory*. American Psychological Association, Washington D.C.: PsycTESTS.
- Swanson, H. L. (2017). Verbal and visual-spatial working memory: What develops over a life span? *Developmental Psychology*, 53, 971-995.
- Swanson, H. L., & Lussier, C. M. (2001). A selective synthesis of the experimental literature on dynamic assessment. *Review of Educational Research*, 71, 321-353.
- Swanson, H. L., & Siegel, L. (2001). Learning disabilities as a working memory deficit. *Issues in Education: Contributions of Educational Psychology*, 7, 1-43.
- Swanson, H. L., Xinhua Zheng., & Jerman, O. (2009). Working memory, short-term memory, and reading disabilities: A selective meta-analysis of the literature. *Journal of Learning Disabilities*, 42, 260-287.
- Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Boston: Pearson.

- Tudge, J., & Scrimsher, S. (2003). Lev S. Vygotsky on education: A cultural-historical, interpersonal, and individual approach to development. In B. J. Zimmerman and D. H. Schunk (Eds.), *Educational psychology: A century of contributions* (pp. 207-228). New York: Routledge.
- Tzuriel, D., & Klein, P. S. (1985). The assessment of analogic thinking modifiability among readers, special education, disadvantaged, and mentally retarded children. *Journal of Abnormal Child Psychology*, *13*, 539-552.
- Unsworth, N., & Spillers, G. J. (2010). Working memory capacity: Attention control, secondary memory, or both? A direct test of the dual component model. *Journal of Memory and Language*, *62*, 392-406.
- Uprichard, S., Kupshik, G., Pine, K., & Fletcher, B. C. (2009). Dynamic assessment of learning ability improves outcome prediction following acquired brain injury. *Brain Injury*, *23*, 278-290.
- Vygotsky, L. S. (1978). Interaction between learning and development. In M. Cole, V. John-Steiner, S. Scribner, & E. Souberman (Eds. And Trans.), *Mind in society: The development of higher psychological processes* (pp. 79-91). Cambridge, MA: Harvard University Press. (Original work published 1935).
- Waber, D.P., Isquith, P.K., Kahn, C.M., Romero, I., Sallan, S.E., & Tarbell, N.J. (1994). Metacognitive factors in the visuospatial skills of long-term survivors of acute lymphoblastic leukemia: An experimental approach to the Rey-Osterrieth Complex Figure Test. *Developmental Neuropsychology*, *10*, 349-367.
- Wilkinson, G. S. (1993). *The Wide Range Achievement Test-3*. Wilmington, Delaware: Wide Range, Inc.
- Wilson, B. A. (2009). *Memory rehabilitation: Integrating theory and practice*. New York: Guilford Press.
- Wilson, K. M., & Swanson, H. L. (2001). Are mathematics disabilities due to a domain-general or a domain-specific working memory deficit? *Journal of Learning Disabilities*, *34*, 237-248.

## Appendix

Table 1

*Characteristics of the Sample*

Frequencies						
	N	Percent of Sample				
<b>Gender</b>						
Females	316	52.49				
Males	286	47.51				
<b>Ethnicity</b>						
White	431	71.59				
Other	171	28.41				
Descriptive Statistics for the Independent and Dependent Variables						
	N	M	SD	d	Minimum	Maximum
Age	602	11.38	3.15	---	4.90	18.00
<b>Digit/Sentence Span-Verbal WM</b>						
Initial	601	2.13	1.42	---	0.00	7.00
Gain	602	3.43	1.90	.91	0.00	9.00
Maintenance	601	2.73	1.77	.44	0.00	9.00
<b>Conceptual Span/Semantic Association (Verbal WM)</b>						
Initial	554	1.40	1.30	---	0.00	7.00
Gain	602	2.57	1.94	.90	0.00	9.00
Maintenance	593	2.06	1.76	.51	0.00	8.00
<b>Visual Matrix (Visual WM)</b>						
Initial	602	2.00	1.23	---	0.00	6.00
Gain	602	4.69	2.06	2.18	0.00	11.00
Maintenance	602	4.00	1.85	1.63	0.00	11.00
<b>Mapping and Direction (Visual WM)</b>						
Initial	601	1.91	1.51	---	0.00	9.00
Gain	602	2.83	1.96	.61	0.00	9.00
Maintenance	602	2.36	1.76	.30	0.00	9.00
<b>WRAT</b>						
Reading	298	49.35	21.24	---	0.00	92.00
Math	298	21.44	11.01	---	0.00	56.00

M=Mean; SD=Standard Deviation; d=Cohen's d. Cohen's d was based on a comparison of initial scores with dynamic assessment scores.

Table 2

*Correlations Between Independent and Dependent Variables*

	1	2	3	4	5	6	7	8	9	10	Reading	Math
1. Ethnicity	---										.31	.18
2. DG (I)	.31	---									.40	.46
3. Verbal (I)	.27	.88	---								.31	.36
4. Visual (I)	.31	.89	.58	---							.40	.46
5. DG (G)	.23	.80	.70	.71	---						.49	.54
6. Verbal (G)	.18	.71	.72	.54	.91	---					.41	.46
7. Visual (G)	.23	.72	.52	.76	.89	.62	---				.47	.51
8. DG (M)	.23	.84	.75	.75	.86	.78	.76	---			.46	.54
9. Verbal (M)	.24	.73	.77	.52	.77	.82	.56	.89	---		.38	.45
10. Visual (M)	.31	.75	.53	.80	.73	.53	.79	.86	.53	---	.43	.49

DG=Domain General; I=Initial; G=Gain; M=Maintenance

p < .001 for all correlations with reading and math

Table 3

*Dynamic Assessment of a Domain General Construct of WM on Reading and Math*

Model 1a (Initial Scores, Domain General WM)				
Achievement	Reading		Math	
Predictors	$\beta$	$F$	$\beta$	$F$
Base Model				
Covariate: Ethnicity	.29**	17.72**	.14*	4.41*
Initial, General Domain	.21**	14.09**	.21**	16.34**
Age: Main Effect	.17**	9.34**	.33**	40.29**
Age x General Initial	-.00	0.01	.08*	4.80*
Total $R^2$	---	.22		.32
Model 1b (Gain Scores, Domain General WM)				
Achievement	Reading		Math	
Predictors	$\beta$	$F$	$\beta$	$F$
Base Model				
Covariate: Ethnicity	.27**	17.19**	.13*	4.95*
DA Measures				
General (G)	.40**	41.56**	.34**	34.46**
Age: Main Effect	.10	3.10	.24**	21.31**
Age x General (G)	-.05	1.68	.11**	7.52**
Total $R^2$		.29		.37
Model 1c (Maintenance Scores, Domain General WM)				
Achievement	Reading		Math	
Predictors	$\beta$	$F$	$\beta$	$F$
Base Model				
Covariate: Ethnicity	.28**	17.87**	.13*	4.75*
DA Measures				
General (M)	.33**	33.14**	.32**	35.65**
Age: Main Effect	.11	3.79	.26**	25.55**
Age x General DA (M)	-.04	1.01	.07	3.07
Total $R^2$		.27		.36
Model 1d (Initial, Gain, and Maintenance Scores, Domain General WM)				
Achievement	Reading		Math	
Predictors	$\beta$	$F$	$\beta$	$F$
Base Model				
Covariate: Ethnicity	.27**	16.45**	.13*	4.11*
Initial, General Domain	-.14	2.71	-.11	2.09
DA Measures				
General (G)	.33**	10.81**	.24*	6.57*
General (M)	.22*	4.32*	.24*	5.80*
Age: Main Effect	.10	2.97	.22**	16.11**
Age x General DA Initial	.16*	4.09*	.10	1.55
Age x General DA (G)	-.06	0.50	.21*	6.24*
Age x General DA (M)	-.15	2.07	-.21*	4.34*
Total $R^2$		.31		.39

DA=Dynamic Assessment; (G)=Gain Score; (M)=Maintenance Score

\* $p < .05$ , \*\* $p < .01$



Table 4

*Dynamic Assessment of WM by Modality on Reading and Math*

Model 2a (Initial Scores, Verbal Domain WM)				
Achievement Predictors	Reading		Math	
	$\beta$	$F$	$\beta$	$F$
Covariate: Ethnicity	.28**	16.20**	.12	3.68
Initial, Verbal Domain	.04	.57	.03	0.36
Initial, Visual Domain	.20**	10.79**	.21**	13.91**
Age: Main Effect	.15	7.61**	.33**	39.45**
Age x Verbal Initial	.08	2.23	.03	0.34
Age x Visual Initial	-.08	2.14	.07	1.86
Total $R^2$		.23		.33
Model 2b (Gain Scores, Verbal and Visual Domain WM)				
Predictors	$\beta$	$F$	$\beta$	$F$
Covariate: Ethnicity	.26**	16.35**	.13*	4.57*
DA Verbal (G)	.19**	8.31**	.16*	6.51*
DA Visual (G)	.24**	13.40**	.22**	11.34**
Age: Main Effect	.09	2.35	.24**	19.69**
Age x Verbal DA (G)	.05	0.62	.10	2.83
Age x Visual DA (G)	-.11	3.11	.02	0.11
Total $R^2$		.29		.37
Model 2c (Maintenance Scores, Verbal and Visual Domain WM)				
Predictors	$\beta$	$F$	$\beta$	$F$
Covariate: Ethnicity	.26**	16.21**	.13*	4.17*
DA Verbal (M)	.15**	7.57**	.17**	10.45**
DA Visual (M)	.24**	16.07**	.20**	13.20**
Age: Main Effect	.09	2.55	.26**	23.26**
Age x Verbal DA (M)	.07	2.07	.08	3.39
Age x Visual DA (M)	.12*	6.22*	-.01	0.07
Total $R^2$		.28		.36
Model 2d (Initial, Gain, and Maintenance Scores, Verbal and Visual Domain WM)				
Predictors	$\beta$	$F$	$\beta$	$F$
Covariate: Ethnicity	.25**	14.42**	.12	3.48
Initial, Verbal Domain	-.13	3.22	-.18*	6.54*
Initial, Visual Domain	-.01	.03	.05	0.49
DA Verbal (G)	.19	3.79	.17	3.75
DA Visual (G)	.15	2.34	.08	0.90
DA Verbal (M)	.13	1.93	.15	3.31
DA Visual (M)	.16	3.19	.14	2.84
Age: Main Effect	.08	1.80	.21**	13.74**
Age x Verbal, Initial	.10	2.08	-.02	.07
Age x Visual, Initial	.11	2.02	.14	3.85
Age x Verbal DA (G)	.02	0.03	.21*	6.06*
Age x Visual DA (G)	-.07	0.76	.04	0.24
Age x Verbal DA (M)	-.04	0.22	-.10	1.70
Age x Visual DA (M)	-.17*	4.10*	-.17*	4.78*
Total $R^2$		.32		.41

DA=Dynamic Assessment; (G)=Gain Score; (M)=Maintenance Score

\* $p < .05$ , \*\* $p < .01$

Table 5

*Descriptive Statistics for the Dependent Variables by Age Group*

Age Group	Y	N	M	SD	Minimum	Maximum
5-7	Reading	62	30.77	20.60	0.00	76.00
	Math	62	12.92	8.19	0.00	29.00
8-9	Reading	81	50.74	16.28	18.00	89.00
	Math	81	18.38	7.52	6.00	37.00
10-11	Reading	52	64.52	18.81	3.00	92.00
	Math	52	26.25	10.47	5.00	52.00
12-13	Reading	33	51.24	16.50	24.00	85.00
	Math	33	22.91	8.03	6.00	37.00
14-17	Reading	70	52.01	19.41	6.00	84.00
	Math	70	28.27	12.00	2.00	56.00

Table 6

*ANCOVA Comparisons of Age x Gain Scores When Predicting Math Achievement*

Tukey's HSD Comparisons													
				Age Group 1 (ages 5-7)		Age Group 2 (ages 8-9)		Age Group 3 (ages 10-11)		Age Group 4 (ages 12-13)		Age Group 5 (ages 14-17)	
				Low	High	Low	High	Low	High	Low	High	Low	High
06	Age Group 1 (ages 5-7)												
	Low	32	-.67	.20	---								
	High	49	-.62	.13	---	---							
	Age Group 2 (ages 8-9)												
	Low	27	-.10	.28	---	---	---						
	High	88	-.28	.10	---	---	---	---					
	Age Group 3 (ages 10-11)												
	Low	69	.04	.34	---	---	---	---	---				
	High	97	.47	.12	<.001	<.001	---	<.001	---	---			
	Age Group 4 (ages 12-13)												
	Low	17	---	---	---	---	---	---	---	---	---		
	High	100	.05	.14	---	<.05	---	---	---	---	---	---	
	Age Group 5 (ages 14-17)												
	Low	10	-.91	.42	---	---	---	---	---	<.05	---	---	---
	High	113	.61	.11	<.001	<.001	---	<.001	---	---	---	<.05	<.05

Table 7

*ANCOVA Comparisons of Age x Maintenance Scores When Predicting Math Achievement*

16

				Tukey's HSD Comparisons									
				Age Group 1 (ages 5-7)		Age Group 2 (ages 8-9)		Age Group 3 (ages 10-11)		Age Group 4 (ages 12-13)		Age Group 5 (ages 14-17)	
	n	LSM	SE	Low	High	Low	High	Low	High	Low	High	Low	High
Age Group 1 (ages 5-7)													
Low	51	-.81	.15	---	---	---	---	---	---	---	---	---	---
High	30	-.54	.15	---	---	---	---	---	---	---	---	---	---
Age Group 2 (ages 8-9)													
Low	53	-.32	.15	---	---	---	---	---	---	---	---	---	---
High	62	-.23	.11	---	---	---	---	---	---	---	---	---	---
Age Group 3 (ages 10-11)													
Low	80	.11	.25	<.05	---	---	---	---	---	---	---	---	---
High	86	.51	.13	<.001	<.001	<.01	<.001	---	---	---	---	---	---
Age Group 4 (ages 12-13)													
Low	22	-.87	.56	---	---	---	---	---	---	---	---	---	---
High	95	.15	.15	<.001	<.05	---	---	---	---	---	---	---	---
Age Group 5 (ages 14-17)													
Low	16	-.74	.27	---	---	---	---	---	<.01	---	---	---	---
High	107	.76	.12	<.001	<.001	<.001	<.001	---	---	---	<.05	<.001	---