2010

Predicting Arithmetic Performance from Age and Executive Function Skills

Andrea Molzhon
Virginia Commonwealth University

PREDICTING ARITHMETIC PERFORMANCE FROM AGE AND EXECUTIVE

FUNCTION SKILLS

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science at Virginia Commonwealth University

By: ANDREA MOLZHON
B.S., Longwood University, 2006

Directors: Geraldine Lotze, Ph.D.
Assistant Professor
Department of Psychology
and
Michelle Ellefson, Ph.D.
University Lecturer in Psychology & Education,
University of Cambridge Faculty of Education, England

Virginia Commonwealth University
Richmond, Virginia
December, 2010

Acknowledgements

First, I would like to thank Dr. Michelle Ellefson, Dr. Elisabeth Blagrove, and Dr. Nick Chater for permission to use their data set [originally collected with support from the British Academy (SG-39180) and the Levehulme Trust (F/215/AY)]. Thank you for trusting me with your work.

A special thank you to my co-chairperson and thesis advisor, Dr. Michelle Ellefson, for the extraordinary amount of time and effort you dedicated to helping me along this process. I am grateful that I had the opportunity to work with and learn from someone with such extensive knowledge and passion for research. Also to my co-chairperson and current advisor, Dr. Geri Lotze, for all the kindness, understanding, and support you have offered me throughout this time. Every time I have doubted myself or my work your kind words have picked me back up and inspired me to keep going; I sincerely thank you.

To the other members of my committee, Dr. Terri Sullivan and Dr. Christopher Chin, thank you for offering your time and helpful advice. Your thoughtful questions and recommendations allowed me to create a product in which I feel comfortable and confident.

To my friends, fellow students, and mother, thank you for having patience with me and for always being there when I needed you. This would not have been possible without your support. Also to Jordan, thank you for always going above and beyond in supporting me throughout this lengthy process, I owe you!

Table of Contents

Page

Acknowledgements ... i

List of Tables ... iv

List of Figures ... vi

Abstract ... vii

Predicting Arithmetic Performance from Age and Executive Function Skills ... 1
- Arithmetic ... 4
  - Development of arithmetic skills. ... 6
  - Neurocognitive components of arithmetic. ... 9
- Executive Functions ... 10
  - Dissociable developments in executive functioning. ... 14
  - Neurocognitive components of executive function. ... 16
  - Models of executive function. ... 18
  - Three components of executive function. ... 21
- Working Memory ... 23
  - Measuring working memory. ... 27
  - Age-related changes in working memory. ... 31
  - Neurocognitive developments in working memory. ... 32
  - Working memory and arithmetic achievement. ... 33
- Inhibition ... 36
  - Measuring inhibition. ... 38
  - Age-related changes in inhibition. ... 39
  - Neurocognitive development of inhibition. ... 42
  - Inhibition and arithmetic achievement. ... 43
- Set Shifting ... 45
  - Measuring set shifting. ... 46
  - Age-related changes in set shifting. ... 49
  - Neurocognitive development of set shifting. ... 51
  - Set shifting and arithmetic achievement. ... 53
- Summary ... 55

Purpose of the Study ... 56
- Hypotheses ... 58

Method ... 59
- Participants ... 59
- Measures ... 61
  - Executive functions. ... 62
    - Inhibition. ... 64
    - Set shifting. ... 69

General cognitive ability. 73
Arithmetic. 74
Procedure 76

Results 77
Description of Data, Screenings and Transformations 77
Zero-Order Correlations 80
Correlations within composites 81
Correlations between age and experimental variables. 88
Correlations between arithmetic and experimental variables. 89
Correlations separated by age group. 89
Regression Analyses 96
Set 1: Two-Step models. 97
Sets 2-7: Five-Step models. 101
Path Analyses 105
Model Testing 106
Composite model. 106
Combined model. 108
Final path models. 109

Discussion 113
Age, General Cognitive Ability, and Arithmetic 114
Age and Executive Functions 117
Executive Functions and Arithmetic 118
Relation to previous literature. 123
Working memory. 123
Inhibition. 125
Set Shifting 127
Limitations to the Study 130
Implications and Directions for the Future 133

List of References 137

Appendices

A. Tic-Tac-Toe 158
B. Stop Signal 159
C. Figure Matching 160
D. Arithmetic Matching 161
E. Missing data calculations 162
F. Hierarchical Regression Models Summary Table: Sets 3-7 163

Vita 167

List of Tables

Page

Table 1. Common Working Memory Tasks Grouped by Appropriate Ages.......................... 28

Table 2. Common Inhibition Tasks Grouped by Appropriate Ages...................................... 40

Table 3. Common Set Shifting Tasks Grouped by Appropriate Ages................................... 47

Table 4. Participant Ages (in Years) by Age Group and Gender 61

Table 5. Cognitive Constructs and Independent Variables (IVs) From Each Measure.......... 62

Table 6. Descriptive Statistics for Unstandardized Variables 78

Table 7. One-Way Analysis of Variance (ANOVA) for Age Group Effects across Standardized Variables 82

Table 8. Tukey's HSD Post-Hoc Test Results of Between-Group Comparisons Across Standardized Measures 83

Table 9. Parametric Zero-Order Correlations for Unstandardized Variables, Grouped by Composite 85

Table 10. Non-Parametric Zero-Order Correlations for Unstandardized Variables, Grouped by Composite 86

Table 11. Descriptive Statistics and Parametric Zero-Order Correlations for Composite and Standardized Variables Included in Path Models 87

Table 12. Non-parametric Correlations for Standardized Variables Across Year-2 Participants.. 90

Table 13. Non-parametric Correlations for Standardized Variables Across Year-5 Participants.. 91

Table 14. Non-parametric Correlations for Standardized Variables Across Year-8 Participants.. 92

Table 15. Non-parametric Correlations for Standardized Variables Across Adult Participants 93

Table 16. MATH Composite Model 1: Regression Analysis Summary for Standardized Variables Predicting Composite Arithmetic Performance.. 98

Table 17. MATH-WRAT Model 1: Regression Analysis Summary for Standardized Variables Predicting MATH-WRAT Performance 99

Table 18. MATH-TSMATH Model 1: Regression Analysis Summary for Standardized Variables Predicting MATH-TSMATH Performance 100

Table 19. Summary of Select Five-Step Hierarchical Regression Analyses for Standardized Variables Predicting MATH Composite, MATH-WRAT, and MATH-TSMATH Scores........ 102

Table 20. Partial and Semi-Partial (Semi-Part) Correlations between Standardized IVs and DVs in the Final Step of Each Regression Model 105

Table 21. Goodness of Fit Indices for Theoretical, Modified, and Final Path Models 109

List of Figures

Page

Figure 1. Approximate Locations of Areas of the Prefrontal Cortex from Lateral (A) and Medial (B) Views 13

Figure 2. Three-Factor Model of Executive Function Adapted from Miyake et al. (2000) ... 20

Figure 3. Anderson's (2002) Conceptual Model of Executive Function 22

Figure 4. The Modal Model of Working Memory 25

Figure 5. The Baddeley and Hitch (1974) Model of Working Memory 26

Figure 6. Hypothesized Path Model 60

Figure 7. Average Performance by Age Group on All Standardized Experimental Variables 81

Figure 8. Final Composite Model with Standardized Coefficients and Residuals 110

Figure 9. Final Combined Model with Standardized Coefficients and Residuals 110

Abstract

PREDICTING ARITHMETIC PERFORMANCE FROM AGE AND EXECUTIVE FUNCTION SKILLS

By Andrea Molzhon, B.S.

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science at Virginia Commonwealth University.

Virginia Commonwealth University, 2010.

Directors: Geraldine Lotze, Ph.D.
Assistant Professor
Department of Psychology
and
Michelle Ellefson, Ph.D.
University Lecturer in Psychology & Education,
University of Cambridge Faculty of Education, England

The learning of mathematics can be a difficult process for many students. Understanding the cognitive components that contribute to arithmetic achievement may illuminate sources of difficulty and inform the development of better teaching and learning practices. Executive functions (EFs) have been implicated in the development of arithmetic skills in early childhood, but less is known about this relation across middle childhood and beyond. The current study included individuals ages 6-7, 9-10, 12-13, and 18+ years and examined the contributions of 3 components of EF, working memory (WM), inhibition, and set shifting (SS), to arithmetic skills in two domains. It was hypothesized that age, general cognitive ability, and EFs would have unique and combined influences on both domains of arithmetic: proficiency and fluency. Results from correlation, regression, and path analyses indicated that WM, inhibition, and SS differentially contributed to arithmetic proficiency and fluency. The implications for education and intervention are discussed.

Predicting Arithmetic Performance from Age and Executive Function Skills

The development of a strong foundation of knowledge and skill in the area of mathematics during childhood and adolescence promotes success both within and beyond the classroom. As young adults advance from the school environment to the workplace, studies have shown that mathematical skills contribute to job-related success apart from the contributions of language skills and intelligence (Paglin & Rufolo, 1990; Rivera-Batiz, 1992). Thus, to provide students with the necessary tools to succeed beyond the school-aged years, cognitive and educational research efforts should focus on informing and devising methods and practices that can be adopted into school curricula for the purpose of promoting competency in mathematics. However, as compared to the large number of studies that have been dedicated to reading and language skill development, relatively few studies have investigated the cognitive components that contribute to the development of mathematic skills. Considering the lifelong importance of mathematical competency and ability, it is necessary to expand upon this area of knowledge and gain a more complete understanding of the mechanisms that contribute to achievement across multiple school-aged groups in order to inform the development of strong educational practices.

Highlighting the need for improvements in our current educational system regarding the teaching of mathematics, the United States government recently announced its concern over the stagnant performance of American students in the areas of mathematics and science (Kuenzi, 2008). For example, the most recent report on the *State of America's Children* indicated that 60 percent of fourth graders and 70 percent of eighth graders fall below grade level in mathematics (Children's Defense Fund, 2008). Moreover, the Trends in International Mathematics and Science Study (Gonzales, Williams, Jocelyn, Roey, Kastberg, & Brenwald,

2008) reported that while the majority of U.S. fourth graders demonstrate that they have basic mathematical knowledge about whole numbers and shapes and can apply this knowledge in general mathematical problem solving, less than half of these students can apply this knowledge to solving multistep word problems or complex numerical or spatial problems. The report also found that 69 percent of U.S. eighth graders cannot apply basic mathematical knowledge to solve complex problems, i.e., problems involving fractions, decimals, negative numbers, units of measurement, and/or probabilities, and 94 percent cannot organize and generalize information to solve novel problems and form conclusions based on data. The 2006 results from the Program for International Student Assessment (Baldi, Jin, Skemer, Green, & Herget, 2007) found that the mathematics scores of 15-year-olds living in the U.S. were lower than the average mathematics scores of students from 31 out of the 57 countries included. Unfortunately, such national and international reports are limited in that they can provide only information on how students are performing and do not provide information on the mechanisms behind such performance rates; thus, studies dedicated to understanding the cognitive components that contribute to mathematics achievement may serve to inform efforts geared towards improving education outcomes.

In terms of the components of cognition that have been found to contribute to mathematics achievement, both domain-general cognitive factors, such as general cognitive ability and executive functions, and domain-specific factors, such as subitizing and language, may differentially contribute to performance across various contexts and age groups (e.g., Espy, McDiarmid, Cwik, Stalets, Hamby, & Senn, 2004; Fuchs, Fuchs, Compton, et al., 2006; Griffin, Case, & Sigler, 1994; Jordan, Levine, & Huttenlocher, 1995; Kroesbergen, Van Luit, Van Lieshout, Van Loosbroek, & Van de Rijt, 2009). Although general cognitive ability and executive functions both are considered to be domain-general cognitive constructs (e.g., Friedman, Miyake, Corley, Young, DeFries, & Hewitt, 2006; Kroesbergen et al., 2009), evidence suggests that the cognitive components of executive function are functionally distinct from the processes related to intelligence and may have a larger influence on school performance across multiple academic domains (Blair, 2006; Welsh, Pennington, & Grossier, 1991). Several studies have found evidence to support the strong influence of executive functions on mathematics achievement, specifically, among preschool-aged children (Bull, Espy, Weibe, 2008; Espy et al., 2004), elementary school children (Bull & Scerif, 2001; Mazzocco & Kover, 2007; St Clair-Thompson & Gathercole, 2006; van der Sluis, de Jong, & van der Leij, 2007), and middle school children (van der Sluis et al., 2007). Moreover, unlike general cognitive ability, the components of executive function undergo substantial age-related developments and are considered to be more dynamic and malleable components of cognition. Therefore, the relative influence of executive functions on school achievement may undergo changes with development. However, no studies to date have examined the degree to which different executive function skills relate to basic arithmetic skills from middle childhood to early adulthood and the degree to which age-related differences affect these relations.

The following literature review begins by defining arithmetic and presenting a summary of the evidence describing the development of arithmetic skills, the neurological components of arithmetic performance, and the cognitive components that contribute to arithmetic performance. Next, the review explores the concept of executive functions, including a description of the history and development of this concept, an introduction to the tripartite model of executive functions (Miyake, Friedman, Emerson, Witzki, Howerter, &

Wager, 2000) that formed the theoretical basis for the current study, an overview of the three components of executive function that are examined in this study, as well as a review of the neurological and behavioral evidence of the typical development of executive function skills. The review summarizes the concepts of working memory, inhibition, and set shifting and describes the different approaches that have been taken to measure these constructs within adult and child populations, neurological and behavioral evidence of typical age-related changes in these constructs, and evidence of the role that each of these constructs plays in arithmetic performance. Measurement approaches are reviewed with a special focus on the measures that are utilized in this study so as to provide the reader with sufficient information pertaining to the way each construct will be viewed in the current project. In addition, studies from neuroscience are reviewed to supplement the developmental literature and to highlight the close association between neurological and behavioral development. This review ends with a description of the purpose of the current study as well as the specific hypotheses that were addressed by the analyses. The intention of this review is to provide the theoretical foundation for the subsequent examination of the relation between executive functions, age-related changes in executive functions, and arithmetic performance across different age groups.

Arithmetic

The field of mathematics is far-reaching; it allows for the measurement of quantities and phenomena, the examination of relations between numbers and/or symbols, the approximation of percentages and probabilities, and much more (Fuchs, Fuchs, Compton, Powell, Seethaler, Capizzi, et al., 2006). Often in psychological and behavioral publications, the terms arithmetic and math or mathematics are used interchangeably, and the distinction between the terms is somewhat arbitrary. In the current study, *mathematics* is used as an umbrella term encompassing all branches of mathematical computation and problem solving, i.e., single-digit and multi-digit calculations, simple and complex computations, word problems, etc., *numeracy* refers to basic knowledge of numbers (or number sense), and *arithmetic* refers to both simple and complex mathematical problem solving involving real, i.e., non-symbolic, numbers (see Bogomolny, 1996). The cognitive components that contribute to proficiency and fluency in both simple and complex forms of arithmetic will be the focus of the current study.

Arithmetic skills have been defined in different ways. In order to achieve in arithmetic, children must demonstrate a sufficient degree of both proficiency and fluency when solving arithmetic problems (Kaye, deWinstanley, Chen, & Bonnefil, 1989). Arithmetic *proficiency* may be viewed as a reflection of one's ability to utilize one's knowledge in arithmetic effectively (see Leach, Coyle, & Cole, 2003), while arithmetic *fluency*, or efficiency (Kaye et al., 1989), integrates proficiency with processing speed and relates to the relative speed with which arithmetic problems are processed and accurately solved (Ramos-Christian, Schleser, & Varn, 2008; Smith-Chant & Lefevre, 2003). As noted by Ramos-Christian et al. (2008), perhaps the most important distinction between proficiency and fluency is that proficiency can be achieved without a true understanding of the processes involved in arithmetical computation (for example, using rote memory to recall facts and steps), while fluency in arithmetic is facilitated by such an understanding. There have been reports of significant individual differences in arithmetic proficiency, for example, in studies comparing children with mathematical difficulties (MD) with typically developing children (e.g., Geary et al., 2000; Hanich, Jordan, Kaplan, & Dick, 2001; Jordan & Hanich, 2000). In addition, Geary (1993) reported significant differences in the speed with which children with MD could solve numerical problems versus their faster typically developing counterparts, and Smith-Chant and Lefevre (2003) found that differences in instructional requirements affected adults' performance on an arithmetic fluency task in different ways, i.e., increased instructional requirements did not affect the performance of adults with high fluency but did affect the performance of adults with low fluency. Thus, individual differences exist for both arithmetic proficiency and arithmetic fluency.

Development of arithmetic skills. Ample evidence has indicated that infants are sensitive to numerical properties and changes in number (e.g., Brannon, 2002; Starkey & Cooper, 1980; Starkey, Spelke, & Gelman, 1990; Van Loosbroek & Smitsman, 1990; Xu & Spelke, 2000; review: Butterworth, 2005; contrary evidence: Feigenson, Carey, & Spelke, 2002). As an example, Wynn (1992) presented several experiments that utilized a looking-time procedure to determine whether infants would demonstrate rudimentary arithmetical abilities. In a series of similar experiments, Wynn presented infants with a puppet-like scene using toy objects and found that four-to-five-month-olds looked longer at the situations involving simple addition and subtraction when these operations were paired with incorrect solutions versus when they were paired with correct solutions. As infants tend to look longer at unexpected versus expected events, Wynn interpreted her results as evidence that five-month-olds are capable of mentally calculating simple arithmetical operations. While such early numeracy skills allow very young children to understand basic properties of numbers and operations, the ability to apply this understanding to perform arithmetic calculations develops later in life.

Although human infants may have the ability to understand simple numerical concepts and arithmetical operations (Wynn, 1992), more complex arithmetic skills and knowledge of arithmetic principles develop and are learned over time (Prather & Alibali, 2009). Beyond infancy, Kaufmann and Dowker (2009) noted that many studies have shown that prior to formal schooling, preschool children typically are able to understand and manipulate quantity as well as perform simple calculations (e.g., Bisanz, Sherman, Rasmussen, & Ho, 2005; Ginsburg, 1977; Hughes, 1986; Jordan, Huttenlocher, & Levine, 1992; Siegler & Booth, 2005). For example, Starkey and Gelman (1982) found that most three-year-olds could solve 2 + 1 and add and subtract one if presented with objects or number words, and most five-year-olds could solve 4 + 2; however, only around half of the five-year-olds were able to solve 2 + 4. By five-to-six years of age, children begin to understand that order does not matter in addition (Carpenter & Moser, 1982). Prior to formal schooling, children's arithmetic skills may be restricted by their use of rudimentary counting strategies, i.e., counting both numbers, rather than more advanced strategies, i.e., counting from the largest number (Butterworth, 1999; Carpenter & Moser, 1982). As children age, they begin to use more effective counting strategies (Carpenter & Moser, 1982) and to count to increasingly higher numbers (Fuson, 1988).

At around age six or seven, children begin to demonstrate an ability to retrieve arithmetical facts from memory – specifically pertaining to multiplication (Butterworth, Marchesini, & Girelli, 2003; for review, see Butterworth, 2005) – and they begin to develop proportional reasoning skills for solving word problems (Van Dooren, De Bock, & Verschaffel, 2010; Van Dooren, De Bock, Hessels, Janssens, & Verschaffel, 2005). Butterworth et al. (2003) found that Italian children ages 6 – 10 performed multiplication problems presented as Larger (number) x Smaller (number) faster than problems presented as

Smaller x Larger; an interesting finding given that the Italian educational systems teaches children Smaller x Larger prior to Larger x Smaller. Their results indicated that while children learned the Smaller x Larger format first, they reorganized their memory to prefer Larger x Smaller after learning this format, suggesting that children mentally organize numbers in specific, universal, ways for arithmetic facts (also, see Butterworth, 2005). Moreover, a longitudinal study of Flemish children from second to eighth grade showed that while early elementary students were capable of correctly responding to proportional word problems, performance continued to improve to sixth grade, and that performance improved most dramatically between third and fifth grade (Van Dooren et al., 2005). Also, from sixth grade to eighth grade, students' ability to distinguish between problems requiring proportional reasoning and non-proportional problems began to improve (whereas prior to sixth grade, students made *more* proportional errors with age), though proportional errors still were present in eighth grade, indicating continued development beyond eighth grade.

Elementary school children typically begin to rely more on long-term memory retrieval processes for solving simple calculations and less on effortful, time-consuming counting strategies (Ashcraft, 1982; Geary, Brown, & Samaranayake, 1991; Kaye et al., 1989). As demonstrated by Lovett (1987), children with disabilities in reading are more likely to differ from their typically developing peers on measures of fluency (reading speed) than measures of proficiency (word recognition), and Ramos-Christian et al. (2008) found that children in a higher stage of cognitive development had greater arithmetic fluency skills than children at a lower cognitive developmental level – though both groups performed with the same level of proficiency. Moreover, Geary et al. (1991) found that typically developing first and second grade students were able to retrieve addition facts from memory significantly faster than they had 10 months prior, though children with MD did not demonstrate any gains in computational speed over the same time period. However, the children with MD in this study did display significant gains in achievement scores. Thus, with age and cognitive development, children typically experience the greatest gains in arithmetic fluency, and while potential gains in proficiency may be consistent across typically developing children and children with learning disabilities, atypical development of arithmetic fluency – rather than proficiency – may serve as the defining characteristic of typical versus atypical development of arithmetic skill (see Bull & Johnston, 1997 and Swanson & Beebe-Frankenberger, 2004).

Neurocognitive components of arithmetic. As evidenced by the studies reviewed thus far, the study of arithmetic knowledge and development has been largely behavioral; only recently have researchers begun to examine the neurological basis for numerical processing and arithmetic performance. Studies involving adults with brain lesions found consistent evidence that the left parietal region of the brain is associated with simple calculations (e.g., Delazer, Karner, Zamarian, Donnemiller, & Benke, 2006; Lee, 2000; Van Harskamp and Cipolotti, 2001; Warrington, 1982) while the frontal area of the brain is associated with complex calculations (e.g., Lucchelli and De Renzi, 1993; Semenza, Miceli, Girelli, 1997; for review, Zamarian, Ischebeck, & Delazer, 2009). Moreover, Zamarian et al. (2009) reviewed several studies that isolated three parietal circuits as being responsible for the processing of numbers: the *intraparietal sulcus* (bilaterally) for representing quantity, approximate computations, and subtraction (e.g., Lee, 2000; Stanescu-Cosson, Pinel, van De Moortele, Le Bihan, Cohen, & Dehaene, 2000), the *angular gyrus* for regulating exact and automated calculations and for retrieving arithmetical facts (e.g., Lee, 2000), and the *superior parietal lobule* for supporting the visuo-spatial processes involved in number processing (e.g., Dehaene, Piazza, Pinel, & Cohen, 2003). Thus, bilateral structures of the parietal region of the brain seem to be associated strongly with number processing, while left parietal regions, specifically, seem to be involved in performing simple calculations – both essential to arithmetic performance.

Though small in number, the developmental studies that have examined the neurological correlates of arithmetical problem solving and number processing, in general, across different age groups have consistently reported that children, in comparison with adults, rely more on prefrontal regions of the brain while solving numerical tasks (Cantlon, Libertus, Pinel, Dehaene, Brannon, & Pelphrey, 2009; Kaufmann, Koppelstaetter, Siedentopf, Haala, Haberlandt, Zimmerhackl, et al., 2006; Kucian, von Aster, Loenneker, Dietrich, & Martin, 2008; for a review, Zamarian et al., 2009). Over time, networks of the brain become more specialized for number processing – likely due to both experience and maturation – and patterns of brain activation associated with arithmetical processes shift from the general prefrontal region to localized number processing centers in the parietal region (Rivera, Reiss, Eckert, & Menon, 2005). The age-related shift from general to specific and from globalized to localized neural activity has been documented in studies examining age-related differences in other areas of neuro-cognitive development, i.e., executive functions, as well (e.g., Crone, Donohue, Honomichl, Wendelken, & Bunge, 2006b; Durston, Thomas, Yang, Uluğ, Zimmerman, & Casey, 2002; Tamm, Menon, & Reiss, 2002).

Executive Functions

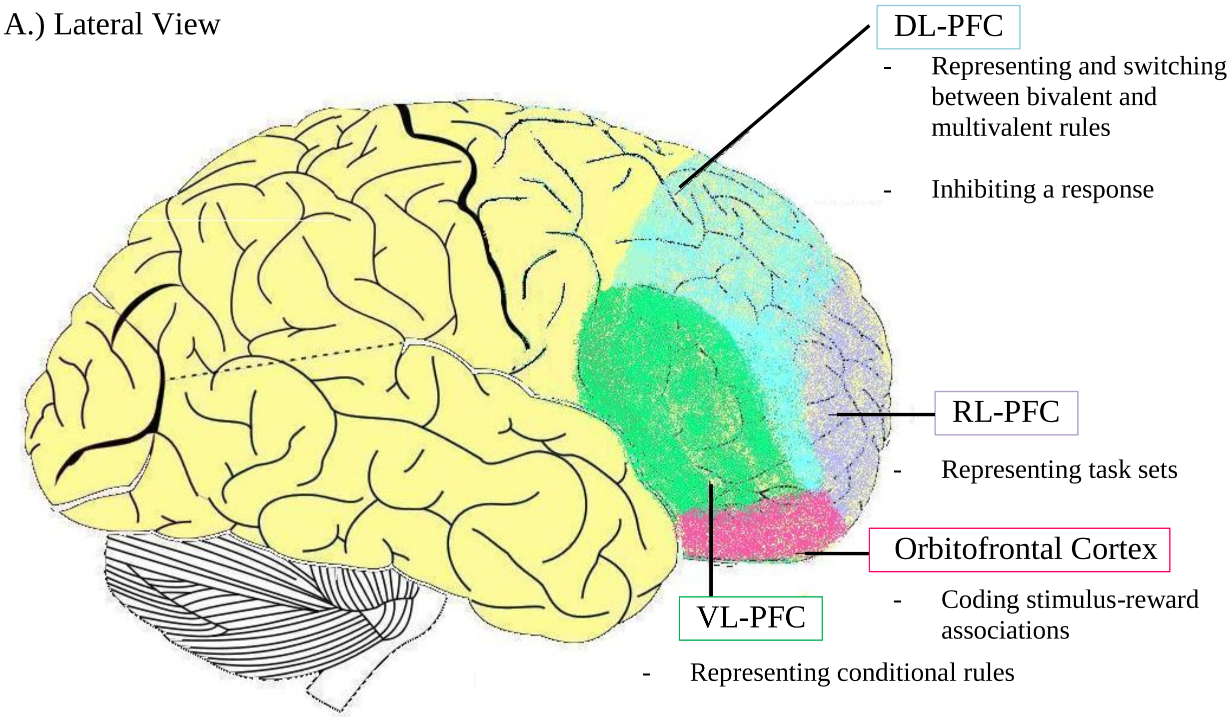
In the field of cognitive development, the term executive function has yet to be defined in clear and universally accepted terms. However, most would agree that executive functions include controlled cognitive processes that are implemented under cognitively difficult circumstances, typically for the purpose of achieving a goal or accomplishing a task (see Baddeley, 1996, Pennington, 1997 and Welsh, Friedman, & Spieker, 2006 for reviews). Interest in this area of cognition largely developed from earlier studies involving individuals with frontal lobe damage. Across multiple studies, researchers consistently found that adults with damage to their frontal lobe, also known as the prefrontal cortex (PFC), tended to display deficits in the same central areas of cognition (e.g., Luria, 1966; Shallice, 1982). Specifically, these individuals showed impairments in the areas that we now consider to be executive functions, such as: planning, goal orientation, cognitive and behavioral inhibition, rule representation and maintenance, and cognitive flexibility (Welsh & Pennington, 1988). The studies that emerged from the fields of neuroscience and cognitive science not only illustrated the functional capacity of the human frontal lobe, but also illuminated the differential areas of cognition that contribute to cognitive control. Thus, the term "executive" refers to the managerial characteristics of the control functions that often are associated with the frontal lobe.

One theory that developed out of studies involving patients with frontal lobe damage was the notion that the neural deficiencies underlying executive function deficits in frontal lobe patients also accounted for the executive deficits observed in children (Kirk & Kelly, 1986). The original theoretical standpoint was that prefrontally-guided executive abilities are essentially non-existent in young children due to the prematurely developed prefrontal cortex (e.g., Golden, 1981). According to this view, damage or lesions to the prefrontal cortex during adulthood causes the adult brain to return to a structural and functional state that is similar to that of a child. However, since the 1980s, studies in developmental psychology have generated considerable evidence against the misconception that the prefrontal cortex does not begin to function until adolescence (e.g., Diamond, 1988; Diamond & Goldman-Rakic, 1985, Welsh et al., 1991; Welsh & Pennington, 1988).

A primary concern for the proposed connection between the cognitive functioning of brain-damaged adults and that of children is that young children show signs of prefrontal function very early in life (Diamond, 1988; Diamond & Goldman-Rakic, 1985; Welsh et al., 1991). Through the course of development, the areas of the PFC become more specialized in function (see Figure 1) and executive skills improve; thus, damage to these specialized areas later in life can lead to loss of function associated with the area of damage. In children, although these areas are still maturing, they are somewhat accessible and able to function on a rudimentary level. Moreover, studies have shown that the underlying causes of behavioral impairments demonstrated by adults with brain damage and individuals with neurodevelopmental disorders are not always the same. Thus, the functional and structural capacity of the PFC in frontal lobe damaged adults is not equivalent to the developing frontal lobe of the young child (see Karmiloff-Smith, 2009; Thomas & Karmiloff-Smith, 2002, for additional evidence that neurodevelopmental disorders are different from acquired brain damage).

In addition to studies examining patients with frontal lobe damage, interest in the area of executive functions emerged out of a prominent theory of working memory that was introduced in the early 1970's. Predating the term "executive function," the concept of a "central executive" component of cognition was proposed in Baddeley and Hitch's (1974) working memory model - a three-component working memory model defining the functional characteristics of the short-term or "working" memory system. In the original model, the "central executive" component of working memory served to integrate and manage the

A.) Lateral View



B.) Medial View

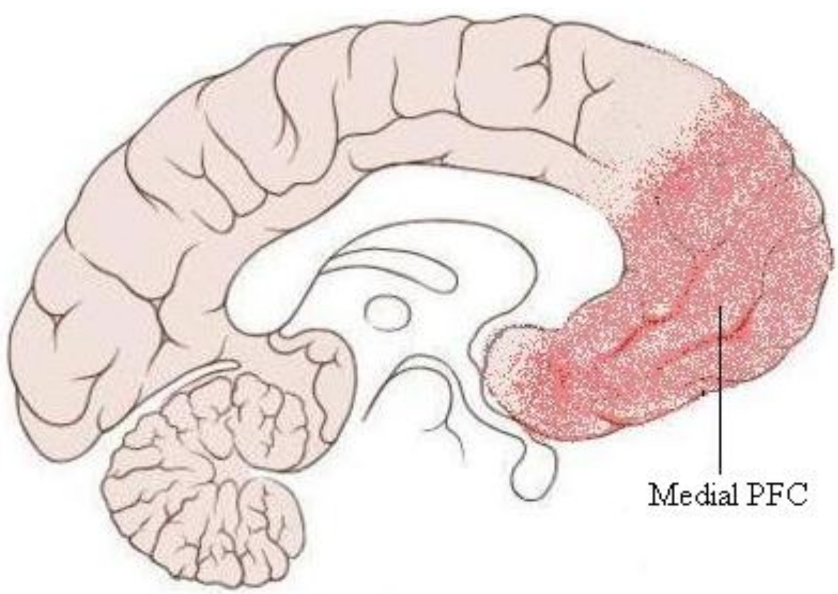


Figure 1. Approximate locations of areas of the prefrontal cortex from lateral (A) and medial (B) views (modified versions of non-copyrighted images obtained through free public license; shading was based on Figure 1 of Bunge & Zelazo, 2006).

information received from the verbal system (the “phonological loop”) and the visual system (the “visuo-spatial sketchpad”), was responsible for the shifting of mental sets, and allowed for the strategic control and direction of attention and inhibition.

More than two decades later, Baddeley (1996) expanded on the original unified definition of the central executive and suggested that the construct is more likely to be a reflection of independent but related cognitive control processes, consistent with the modern diversity theories of executive function. In reconceptualizing the central executive, he rejected the popular assumption that executive control processes are a reflection of activity in the frontal lobe alone. Though he did not dispute neurological evidence that isolated the role of the frontal lobe in tasks of executive control (e.g., Duncan, Emslie, Williams, Johnson, & Freer, 1996), he advised against assuming that executive control is dependent strictly on the functioning of the frontal lobe. Subsequently, support for the hypothesis that executive functions are influenced by brain structures outside of the frontal lobe has arisen from studies examining the functions of the parietal lobe (Alvarez & Emory, 2006; Collette, Van der Linden, Laureys, Delfiore, Degueldre, Luxen, et al., 2005) and the basal ganglia (Alvarez & Emory, 2006; Aron, 2008).

Dissociable developments in executive functioning. Like Baddeley (1996), many researchers have begun to adopt a diverse explanation of executive function, describing this concept as an integration of complex cognitive processes rather than as a singular cognitive construct (Baddeley, 2002; Lehto, Juuja¨rvi, Kooistra, & Pulkkinen, 2003; Miyake et al., 2000). Under this diversity perspective, unique executive processes each play a different role in actively controlling cognition in order to elicit a desired response. Although debates persist concerning the exact role of each of these components during a given executive task, it has been suggested that executive functions include such components as inhibition, set-shifting/switching, working memory/updating, goal/task selection, rule representation, and controlled attention (Bunge & Zelazo, 2006; Diamond, 2002; Kray, Li, & Lindenberger, 2002; Mayr, 2002).

In addition to viewing these processes as unique in function, many researchers have asserted that they develop along unique and distinct pathways (see Anderson, 2002 for a review). In particular, studies examining age-related changes in cognition have found ample evidence favoring the distinctiveness of working memory updating, inhibition, and set shifting from young childhood through adulthood (Brocki & Bohlin, 2004; Bull & Scerif, 2001; Huizinga et al., 2006). Across multiple studies, some general themes have emerged regarding the diverse developmental course of these executive functions. For example, studies have found that three-year-olds have great difficulty inhibiting a strong response, frequently continue to perseverate under conditions of a previously used rule, have extreme difficulty representing bi-dimensionality, and have a weak ability to delay gratification (e.g., Diamond, 2002; Diamond et al., 2002; Diamond & Taylor, 1996; Gerstadt, Hong, & Diamond, 1994; Gopnick & Rosati, 2001; Passler, Isaac, & Hynd, 1985; Zelazo, Reznick, & Piñon, 1995). In contrast, five-year-olds tend to be much more capable of performing these actions with few mistakes, and six-seven-year-olds can perform these basic executive functions as efficiently as adults, i.e., they respond more quickly and make significantly fewer mistakes than their younger counterparts (e.g., Diamond, 2002; Zelazo, Müller, Frye, & Marcovitch, 2003).

In support of a diverse – rather than unified – perspective of executive function, studies have found that performance on more complex executive tasks tends to improve at a slower rate than performance on more basic tasks. For example, children's performance on tasks that require switching between rules, suppressing a dominant response, and holding and manipulating multiple pieces of information in mind may not peak until around 11 years of age, and—depending on the task—sometimes does not reach adult level until 15 or 16 years of age (Diamond, 2002; Huizinga et al., 2006). Moreover, others have found developmental differences in the ability to represent and shift among multifaceted rules versus the ability to shift between rules. Specifically, evidence has shown that children develop the ability to switch between dichotomous rules earlier than they are able to represent complex, conditional rules (Crone et al., 2006b). Processes related to working memory appear to develop earlier than processes related to inhibition and set shifting, and the rate of improvement across each of these components tends to depend upon task complexity (Crone, Bunge, van der Molen, & Ridderinkhof, 2006a; Diamond, 2002; Huizinga et al., 2006). Taken together, these studies lend credence to the diversity perspective, indicating that the processes related to executive function display differential patterns of development, and suggest the importance of considering developmental level when measuring and conceptualizing the components of executive function.

Neurocognitive components of executive function. Adding to the developmental literature, studies in cognitive neuroscience have examined the relation between brain maturation and developments in executive processing from a physiological perspective. Such studies have been made possible in the recent years by improvements in technology that have allowed cognitive neuroscientists to gain a better understanding of the patterns of brain function that are associated with performing certain cognitive functions. Using functional magnetic resonance imaging (fMRI), electroencephalogram (EEG), and event-related potentials (ERPs), researchers have found significant links between patterns of activity in specific areas of the brain and task performance associated with particular executive functions (Dustman, Emmerson, & Shearer, 1996). As previously indicated, the most notable area of the brain that has been shown to relate to components of executive function is the PFC (refer back to Figure 1). Specifically, the orbitofrontal cortex and the lateral PFC, comprised of the ventrolateral PFC (VL-PFC), the dorsolateral PFC (DL-PFC), and the rostrolateral PFC (RL-PFC), typically are activated when one is utilizing executive skills (Bunge & Zelazo, 2006).

Bunge and Zelazo (2006) compiled a comprehensive review of the neurological data indicating that different areas of the PFC are involved in different aspects of rule use over the course of development. Based on the researchers' theoretical perspective, they chose to define developments in executive functioning as a function of increasing levels of rule complexity. As rule complexity increases, patterns of brain activation tend to shift from one area of the PFC to another. Specifically, the orbitofrontal cortex is implemented when representing a single rule, VL-PFC and DL-PFC are related to representing bivalent and multivalent rules, and the RL-PFC is related to representing task sets. While representing rules that are associated with a common stimulus, i.e., *conditional rules,* is most associated with VL-PFC activity, switching between two bivalent rules, i.e., switching from color (red or blue) to shape (circle or square), is associated with activation of the DL-PFC. A possible explanation for the differential patterns of PFC activation associated with different executive demands is that the VL-PFC may be more highly related to basic rule representation while the DL-PFC is related to inhibiting a previously used rule. One of the studies cited in this review examined patterns of electrophysiological activity in nonhuman primates and found evidence indicating that the orbitofrontal cortex is largely related to encoding the association between a stimulus and a reward (Wallis & Miller, 2003). Moreover, evidence has shown that the orbitofrontal area of the PFC matures dramatically over the first three years of life, coinciding with major improvements in the ability to mentally reverse stimulus-reward associations (Overman & Bachevalier, 1999).

What is perhaps most interesting about the neurological components of executive functioning is that these components seem to mature and change at the same rate as executive function development (Aron, 2008; Bunge & Zelazo, 2006; Diamond, 2002). The number of neuronal connections in each area of the PFC generally reaches adult level in a specific order: first in the orbitofrontal PFC, then in the VL-PFC, and finally in both the DL-PFC and the RL-PFC (Bunge & Zelazo, 2006). Interestingly, improvements in rule use tend to follow the same developmental course as the rate of maturation in the lateral PFC. This shared course of development lends further support to the theory that age-related developments in executive functioning are related to the maturation of the PFC and to age-related refinements in several neural networks connected to the PFC. In addition, the results from these studies have supported the functional separation of the components of executive function and have offered further support to the notion of unique developmental trajectories associated with each distinct component.

Models of executive function. Rather than viewing executive function from within the confines of the working memory model, most of the recent literature has conceptualized executive function as a cognitive control system largely independent from other models of cognition. This way of conceptualizing executive function has led to the development of several multi-component theoretical models – models that ascribe to the diversity perspective that was previously described. Typically, these models are derived by way of factor or latent variable analysis to identify the common and distinct structural characteristics of various measures of executive function. For example, Miyake and his colleagues (2000) used latent variable analysis to examine the underlying components measured by a standard battery of executive tasks. The results of their analysis lead to the development of the three-component model that, perhaps, represents the most highly accepted model of executive function to date. Their model identifies working memory updating, inhibition, and set shifting as the three latent components that comprise the concept of executive functions (see Figure 2). As Miyake et al.'s is the most accepted model of executive function, the next sections will describe these components in more detail as they are critical to the theoretical basis of the present study.

Prior to the development of specific models of executive function with adults, Levin and his colleagues (Levin, Culhane, Hartmann, Evankovich, Mattson et al., 1991) tested children on a battery of "frontal lobe" tasks and used principal components analysis to group the variables from each of the tasks into one of three component constructs. Levin et al.'s approach differs from later approaches to modeling executive function in that he chose to combine common measures of executive functions, i.e., an inhibition task, a sorting task, and a planning task, with other purported measures of frontal lobe functioning, i.e., verbal learning tasks, in developing three integrated constructs of frontal lobe function: *semantic association and concept formation*, *freedom from perseveration*, and *planning and strategy* (Levin et al., 1991). In addition, unlike Miyake et al.'s (2000) model of adult executive functions, Levin et al.'s component model was based on a developmental study that involved children ranging in age from 7 to 15 years. Generally, the results from this study are

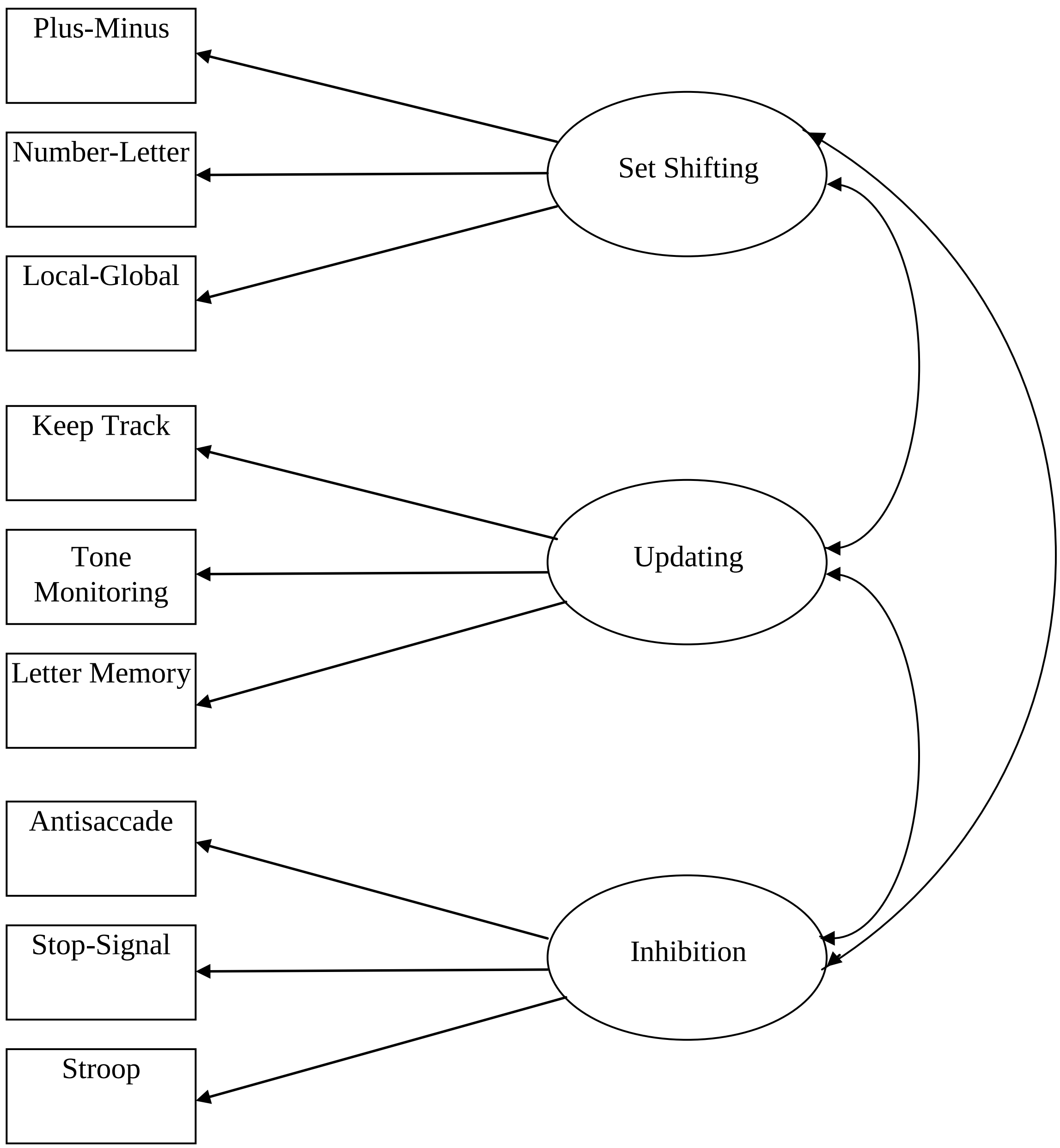


Figure 2. Three-factor model of executive function adapted from Miyake et al. (2000). Single-headed arrows represent regression paths and double-headed curved arrows represent correlations. Boxes represent variables that were measured directly and circles represent the latent variables that were identified through factor analysis.

consistent with the findings of several subsequent developmental studies that have specifically examined age-related changes in executive functioning (e.g., Anderson, Anderson, Northam, Jacobs, & Catroppa, 2001; Huizinga & van der Molen, 2007; Welsh et al., 1991). For example, like Levin et al., Anderson et al. (2001) found that planning performance continued to improve beyond late childhood and into adolescence. In addition, consistent with Levin et al., studies conducted by Huizinga and van der Molen (2007) and Welsh et al. (1991) found that cognitive flexibility improved until around 11 years of age. Although the three-component model has amassed the most empirical and theoretical support in the recent literature, alternate multi-component models of executive function exist within the literature, as well. Consistent with the latent variable approach utilized by Miyake et al. (2000) and the developmental approach of Levin (1991), Anderson (2002) suggested that developments in executive function occur in the areas of cognitive flexibility, attentional control, goal setting, and information processing (see Figure 3). Like Miyake et al. (2000) and Welsh et al. (2006), Anderson proposed that the potential components of executive function are distinct but related, and that performing an executive task requires the *coordination* of these components. Although the developmental models proposed by Levin (1991) and Anderson (2002) do not isolate specifically the three components of working memory updating, inhibition, and set shifting, recent studies have found that these three components are present in children as young as six, and that the components of executive function are distinct from other frontal lobe functions across development (Huizinga et al., 2006; Huizinga & van der Molen, 2007).

Three components of executive function. Over the past decade, the majority of studies that have examined executive function or executive function development have

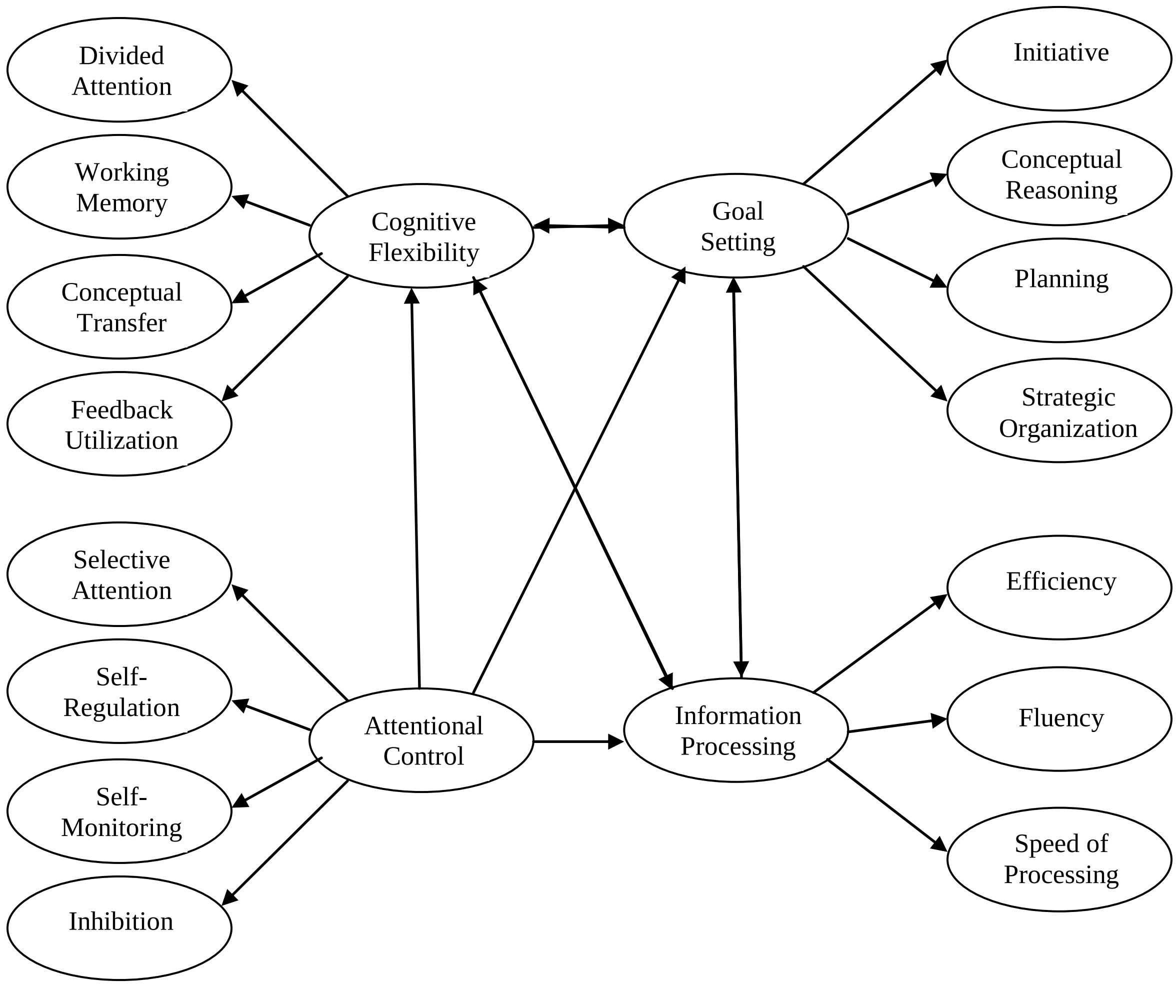


Figure 3. Anderson's (2002) conceptual model of executive function.

focused on at least one of the components described in Miyake et al.'s (2000) three-component model: working memory updating, inhibition, and set shifting. Since the publication of their model, subsequent studies have supported the claim that working memory updating, inhibition, and set shifting represent the three most essential components of executive function (e.g., Huizinga et al., 2006; Welsh et al., 2006). As an example, at the start of 2010 – a decade since the model first was introduced – a popular search database indicated that the Miyake et al. publication had been cited in at least 766 subsequent publications (PsychInfo®, 2010). Thus, in the present review, the definitions of working memory updating, inhibition, and set shifting have been adopted from the operational definitions that were proposed by Miyake and his colleagues (2000).

In the context of three-component model, updating refers to the facet of working memory that allows one to actively update and recall information that is presented during a given task. Inhibition, or inhibitory control, also is necessary in order to perform tasks that require cognitive control as it pertains to one's ability to actively suppress a dominant, habitual, or previously used response in order to satisfy a given rule or condition. Set shifting relates to one's ability to flexibly and efficiently alternate between different response patterns, synonymous with the terms switching and cognitive flexibility. Although Miyake et al. (2000) found that the three-factor latent variable model was the most empirically and theoretically supported by their data, their results indicated that the three components were correlated moderately and not completely separable. Thus, the authors concluded that the component processes related to executive function are separable and distinct, though they share a common basis. As previously indicated, the coming sections of this review will focus on each of these components individually. The core ideas that will be highlighted in each of these sections will pertain to forms of measurement, age-related developments, neurocognitive developments, and relation to arithmetic achievement. The literature review will conclude with summary of the evidence for the relation between executive functions (working memory updating, inhibition, and set shifting) and arithmetic achievement.

Working Memory

Working memory, in general, represents the processes of the memory system that are responsible for the active updating, manipulation, storage, and retrieval of incoming information. Although the concept did not emerge in the literature until the turn of the twentieth century, psychologists already had begun to distinguish between the components of memory by the late nineteenth century (e.g., James, 1890/1950). Predating the concepts of "short-term," "long-term," and "working memory," the idea that memory could be divided into a primary component and a secondary component was proposed by William James in 1890 and was expanded upon by Waugh and Norman in 1965 (Cowan, 2005). The term *primary memory* refers to immediate and temporary storage of information and is most akin to the concept of short-term memory or working memory; *secondary memory* is nearly identical to the concept of long-term memory and refers to the storage of information across one's lifetime. Following from early theories of primary memory, Miller et al. proposed the concept of working memory and defined it as the process responsible for the short-term maintenance, evaluation, and execution of goal-directed behavior (Miller, Galanter, & Pribram, 1960). In the following decade, many researchers became interested in studying the concept of working memory from an information processing perspective, and this interest lead to the development of the theories and models of working memory that sare accepted today (Cowan, 2005).

As depicted in Figure 4, early theories of working memory, i.e., prior to 1970, share many of the same tenets and fit within the loosely defined "modal model," first sketched by Broadbent in 1958 (termed by Baddeley, 1986). This model has been criticized for two primary reasons: 1) the assumption that unattended sensory information becomes lost forever, and 2) the assumption that only an indirect link exists between the sensory store and the long-term memory store (Cowan, 2005). On the basis of the need for a less simplistic and more accurate model of working memory, Baddeley and Hitch (1974) developed a

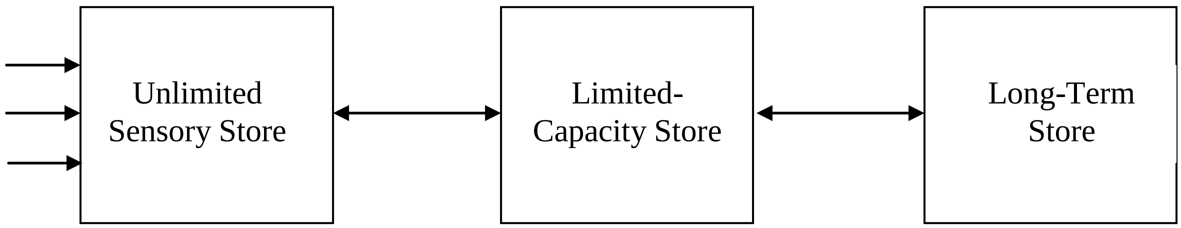


Figure 4. The modal model of working memory (created by Broadbent, 1958; adapted from Cowan, 2005).

multifaceted model of working memory that defined working memory as an integrated system with distinct processing components (see Figure 5). Recall that the original Baddeley and Hitch (1974) model assumed working memory to be comprised of two essential slave systems: the phonological loop and the visuo-spatial sketchpad, as well as one control or managerial system termed the central executive. While the phonological loop is responsible for recognizing, ordering, and briefly storing verbal sounds heard in the environment, the visuo-spatial sketchpad is responsible for recognizing and briefly storing visual information. In terms of the working memory model, the central executive is defined as an attentional control system that coordinates information received from the two slave systems and controls the various cognitive processes involved in the storage, retrieval, and manipulation of information.

Since the debut of the original model, Baddeley has published various reports that have identified and attempted to alleviate some of the flaws he noted in the model. Most notably, he has since added a third slave system to the model – the episodic buffer (Baddeley, 2000). The primary function of the episodic buffer is to organize information received from the visuo-spatial sketchpad and the phonological loop in terms of each event’s approximate

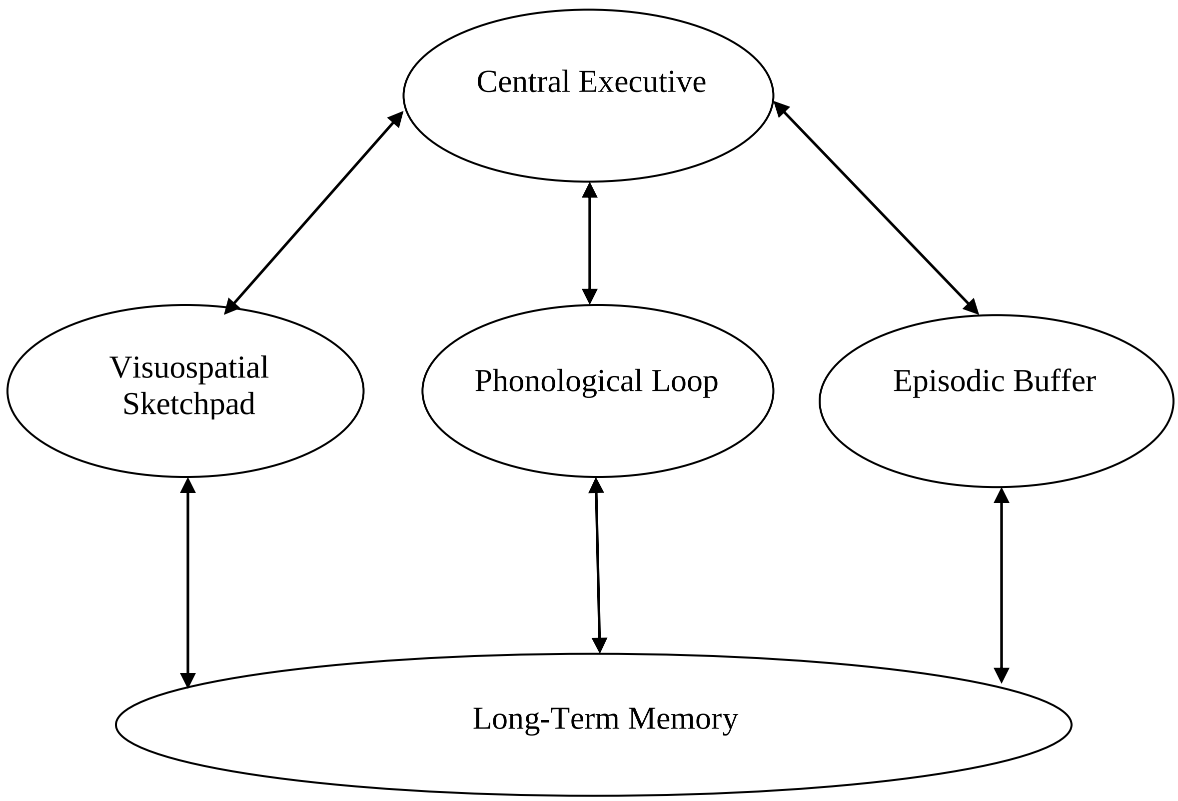


Figure 5. The Baddeley and Hitch (1974) model of working memory with the addition of the episodic buffer (Baddeley, 2000) and the explicit relation between the storage buffers and long-term memory (adapted from Cowan, 2005).

chronological sequence. In addition, he expanded on the original concept of the central executive in his publication, *Exploring the Central Executive* (Baddeley, 1996). In this article, he renounced the unitary definition of the central executive in favor of the now popular diversity perspective – the view that assumes the central executive to be comprised of a variety of separate but related processes.

Over the past decade, researchers have become increasingly interested in the relation or potential overlap between working memory processes and executive function. While some researchers have assumed executive functions to be part of the working memory system's central executive – with working memory underlying all executive functions (Baddeley, 1996; Baddeley & Hitch, 1974; De Rammelaere, Stuyven, & Vandierendonck, 2001; DeStefano & LeFevre, 2004; Luciana & Nelson, 1998; Oberauer, Süß, Schulze, Wilhelm, & Wittmann, 2000; Swanson & Sachse-Lee, 2001), others have chosen to include only a portion of working memory function in their models of executive function (Friedman et al., 2007; Huizinga et al., 2006; Miyake et al., 2000; Welsh et al., 2006). For example, in their theoretical model, Miyake et al. (2000) chose to include only one component of working memory, i.e., updating, as well as one component for inhibition and one component for set shifting. On the other hand, researchers such as Swanson, Jerman, and Zheng (2008) and De Smedt, Janssen, Bouwens, Verschaffel, Boets, and Ghesquière (2009) have included multiple aspects of working memory (e.g., working memory span, updating, and semantic association) in their research design in order to tap the processes of the so-called central executive. Thus, what working memory researchers view to be measures of the central executive and what executive function researchers view to be measures of the executive component of working memory may be one in the same, and working memory studies have not provided the evidence necessary to support the claim that working memory underlies all forms of executive function.

Measuring working memory. Like other components of executive function, many tasks have been designed to measure some form of working memory; Table 1 provides examples of such tasks. The task batteries that different researchers have chosen to use throughout the literature have varied depending on the researcher's theoretical perspective, developmental considerations, and/or research goals. Oberauer et al. (2000) performed a

Table 1.

Common Working Memory Tasks Grouped by Appropriate Ages

Task	Reference	Task Demands
Child tasks		
Bead Memory	Thorndike, Hagen, & Sattler (1986)	Find bead that matches experimenter's bead (after delay)
		Form pattern of beads that matches experimenter's pattern
Spin the Pots	Hughes & Ensor (2005)	Remember multiple locations where objects are hidden
		Retrieve objects after locations have been covered and rotated
Counting and Labeling	Gordon & Olson (1998)	Listen to experimenter label, count, and label & count objects
		Do the same for new set of objects
Older child ($\geq$ 6 years) & adult tasks		
Forward Digit Span[a,b]	Elliot (1996)	Recall strings of digits
Backward Digit Span[b]	Elliot (1996)	Recall string of digits in reverse order
Computation Span	Oberauer et al. (2000)	Indicate whether solutions provided for simple equations are true/false
		Recall solutions from each equation presented per series
Reading Span	Daneman & Carpenter (1980)	Read series of sentences
		Perform distractor task
		Recall last word from each sentence

Task	Reference	Task Demands
Verbal Span	Oberauer et al. (2000)	Memorize list of words
		Perform distractor task
		Recall list of words
Letter Memory	Morris & Jones (1990)	Series of letters shown
		Name last 4 letters from each series
Memory Updating – Figural	Salthouse, Babcock, & Shaw (1991)	Mentally rotate dot patterns several times
		Indicate new location of dots
Spatial Working Memory	Oberauer et al. (2000)	Mentally rotate series of patterns
		Draw all rotated patterns in series
Tic-Tac-Toe[a]	Huizinga et al. (2006)	Remember visual pattern
		Respond when all elements of pattern have been displayed
Tone Monitoring	Larson, Merritt, & Williams (1988)	Hear tones of different pitches
		Respond to 4th tone of each pitch
Keep Track	Yntema (1963)	Listen to/read series of words
		Recall words from given category

[a]Tasks that will be included in the present study. [b]May be used with children ≤ 6 years.

complex factor analysis on adult data sampled from 23 working memory tasks used within the literature and grouped these tasks into one of three content areas: verbal, numerical, or figural-spatial. Vock and Holling (2008) selected six complex-span tasks from the original Oberauer et al. (2000) battery of 23 tasks that were most likely to represent each facet of working memory, had the strongest psychometric properties, and were appropriate for use with children. The tasks used in Vock and Holling's (2008) test battery is provided in Table 1, along with those included in the Miyake et al. (2000) test battery (as their study provides part of the theoretical basis for the current study), those utilized in the present study, and several other working memory tasks that have been used with children.

According to Oberauer et al. (2000), different tasks tap different functional components of working memory. They defined these functional areas as storage and transformation, supervision, and coordination. Although their theory suggests that the supervision function of working memory represents executive function, as previously indicated, most theories of executive function assume that executive processes such as inhibition, cognitive flexibility, planning, and goal-directed behavior are related to the working memory system but are not controlled by this system. Despite theoretical differences, the storage and transformation function of Oberauer et al.'s (2000) theory is highly similar to the measures of working memory that are common in studies of executive function. Moreover, studies that have examined working memory as an executive function have tended to focus on the updating component of the storage and transformation function of working memory. Included in their three-factor theory of executive functions, Miyake and his colleagues (2000) defined updating as the active process by which incoming information is monitored and updated. While other working memory processes related to storage and transformation, such as rehearsal and retrieval, undoubtedly require a degree of executive control, updating is the component of working memory that has been most frequently linked to performance on traditional tasks of executive function (Miyake et al., 2000; Huizinga et al., 2006; van der Sluis et al., 2007). The present study includes two working memory span tasks, forward digit span (Elliot, 1990) and the Tic-Tac-Toe task (Huizinga et al., 2006, adapted from Milner, 1971), that reflect the updating component of working memory.

Age-related changes in working memory. The development of the ability to actively update and maintain information in working memory has been studied extensively in both humans and primates (Diamond, 1990; Espy & Kaufmann, 2002; McGuigan, & Núñez, 2006; Munakata, 1998). Studies have found that the processes related to working memory and inhibition display early signs of development within the first two years of life and continue to improve and refine throughout childhood (Diamond, 1990; Diamond, 2002). In addition, Gathercole, Pickering, Ambridge, and Wearing (2004a) found that the three original components of the Baddeley and Hitch (1974) working memory model are present and functioning in children as young as six, and that the capacity of each of these components continually improves from age four until early adolescence.

To investigate developments in executive working memory over the course of adolescence, Luciana, Conklin, Hooper, and Yarger (2005) compared the performance of individuals from five age groups (ages 9-10, 11-12, 13-15, 16-17, and 18-20) on multiple measures of working memory, ranging in degree of executive control required. They defined executive working memory as the cognitive process that is required when one faces a delay between the presentation of information and the ability to respond to that information, when one must hold multiple items in mind in a fixed temporal sequence, and/or when one must actively organize information in mind in the absence of external cues (Luciana et al., 2005). Based on participants' performance on a delayed response task requiring spatial working memory, the authors found that this aspect of executive working memory showed the most signs of improvement from ages 16 to 20, with marginal improvements detected after age 13. In contrast, results from a self-ordered search task, in which participants were required to recall and strategically order information held in working memory, indicated that this aspect of executive working memory develops most from 9 to 16 years of age and stabilizes at around 18 years of age. In addition, in a cross-sectional study involving 7-year-olds, 11-year-olds, 15-year-olds, and 21-year-olds, Huizinga et al. (2006) found that on two out of three working memory tasks, adult-level performance was not reached until 15 years of age. Thus, the results from these studies not only support the claim that executive working memory continues to develop beyond childhood, but also that this component of executive function develops differentially depending on the executive demands of the task.

Neurocognitive developments in working memory. As previously indicated, studies that have emerged from the combined fields of cognitive neuroscience and developmental psychology have found clear links between specific maturations in the brain and the developmental time course of many executive processes (Bunge & Zelazo, 2006; Diamond, 2002; Huttenlocher & Dabholkar, 1997). Diamond (1985; 1990; 2002) has found that age-related differences in performance on typical tasks of working memory and inhibition increases steadily over the first year of life at the same rate that maturational changes occurs in DL-PFC and connections between the parietal and frontal lobes begin to strengthen (refer to Figure 1). Although the ability to store and update information in working memory develops early in life, the amount and quality of information that can be held in working memory increases with development (Luna, Padmanhaman, & O'Hearn, 2010). For example, Geier, Garver, Terwilliger, and Luna (2009) found that older children, adolescents, and adults all recruited areas of the parietal lobe and frontal lobe when performing a working memory task while undergoing fMRI; however, they found that the children and adolescents relied on the DL-PFC during delay periods more heavily than adults who relied more on parietal regions and the IFC. The authors concluded that adults are more likely to use more specialized neural regions when utilizing working memory than are children and adolescents. Thus, evidence of age-related neurological differences in working memory processing, along with behavioral reports (e.g., Huizinga et al., 2006), indicate that working memory processes may continue to develop through childhood and into adolescence.

Working memory and arithmetic achievement. Considering the diverse functional capabilities of working memory as well as the relatively early developments that typically occur in working memory system, it is of no surprise that working memory has been linked to school achievement starting from as early as preschool. However, much of the literature on the relation between school achievement and working memory has conflicted in terms of theoretical perspective, experimental measures used, and conclusions drawn. For example, Gathercole and Pickering (2000) concluded that the central executive component of the working memory system provided the strongest predictor of literacy and arithmetic performance at seven years of age and again at eight years of age. In contrast to studies of executive function, the three measures used in their study to represent the central executive were recall tasks similar to the backward digit span task summarized in Table 1. In fact, none of the 13 measures used in this test battery of working memory are consistent with the working memory tasks used in Miyake et al's (adult; 2000) or Huizinga et al.'s (developmental; 2006) executive function test batteries. These discrepancies may be due, in part, to the lack of consensus surrounding the definition of the central executive and the role of working memory as an executive function (Espy et al., 2004). Thus, the measured effects of working memory function on school achievement may be dependent on developmental level *as well as* the component of working memory being assessed. Nonetheless, the general conclusion that may be drawn from the wide range of literature that currently exists is that working memory, like other components of executive function, provides differential contributions to school achievement across development.

To this point, researchers interested in the relation between working memory and school achievement have focused primarily on early childhood (Bull et al., 2008; Bull & Johnston, 1997; Espy, et al., 2004) and middle childhood (Adams & Hitch, 1997; Gathercole, Pickering, Knight, & Stegmann, 2004b; McLean & Hitch, 1999; St-Clair-Thompson & Gathercole, 2006); although some have examined this relation in adults, as well (De Rammelaere et al., 2001; Seyler, Kirk, & Ashcraft, 2003). Like other studies in executive function, many of these studies have used performance on memory span tasks (i.e., Bull et al., 2008; Seyler et al., 2003; St-Clair-Thompson & Gathercole, 2006), span and speed tasks (i.e., Adams & Hitch, 1997; Bull & Johnston, 1997; Gathercole et al., 2004), and span and spatial location tasks (i.e., McLean & Hitch, 1999) to operationalize the concept of working memory. While some studies that have examined the relation between working memory and early school achievement have focused on reading and mathematics performance (e.g., St-Clair-Thompson & Gathercole, 2006), other studies have found links to science achievement, as well (e.g., Gathercole et al., 2004b). In general, studies have found that the role of working memory in verbal and mathematical skill development during childhood varies depending on age.

While Bull et al. (2008) operationalized working memory as a reflection of performance on forward and backward span tasks and found that working memory processes contributed to arithmetic proficiency (as measured by an achievement test) in children from preschool through age seven, and St-Clair-Thompson and Gathercole (2006) operationalized working memory as a reflection of performance on a letter memory and a keep-track task (refer to Table 1) and found that working memory uniquely predicted arithmetic and reading proficiency at 11 years of age (measured by standardized achievement test scores). In addition, Gathercole et al. (2004b) defined working memory as a reflection of performance on several recall tasks, a matching task, and a repetition task and found that the relation between working memory and English scores, unlike arithmetic scores (also a measure of proficiency), did not remain constant from 7 to 14 years of age. What seems to be fairly consistent across the literature, however, is the finding that working memory processes are more commonly recruited at younger ages when children are first learning to encode relations between numbers, sounds, and letters; whereas in older children, working memory is recruited under more complex conditions.

Although some general conclusions may be drawn concerning the changing role of working memory in school achievement across childhood, the role of working memory in math achievement, in particular, is slightly more difficult to define under general terms. For example, Trbovich and LeFevre (2003) found that adjusting the format (vertical versus horizontal presentations) of arithmetic problems led to the recruitment of different working memory processes, i.e., vertical presentations related to visual-spatial working memory and horizontal presentations related to phonological working memory. In addition, while some studies have indicated that working memory provides a unique contribution to arithmetic proficiency during the school-age years (Bull & Scerif, 2001), Espy et al. (2004) reported that the contributions of working memory might overlap significantly with the contributions of other executive functions in preschool-aged children. This may be due to a greater need for general executive function skills during the beginning stages of learning versus more case-specific and localized needs for executive function required in more complex learning experiences – experiences that typically occur later in childhood (Espy et al., 2004; DeStefano & LeFevre, 2004). Moreover, studies have shown that older children (beyond around seven years of age) rely on working memory less for simplistic, i.e., single-digit, arithmetic calculations and more for solving complex equations (Ashcraft & Kirk, 2001; Furst & Hitch, 2000), performing mental calculations (Ashcraft, Donley, Halas, and Vakali, 1992; Hitch, 1974; Logie, Gilhooly, & Wynn, 1994), and solving mathematical word problems (Geary, 2004; Lee, Ng, & Ng., 2009; Passolunghi & Siegel, 2001). However, no previous studies examined the relation between working memory and arithmetic fluency (rather than simply arithmetic proficiency); so, at present, no conclusions can be drawn about the relation between working memory and fluency development. Thus, while the current literature indicates that the need for executive control over the storage and updating of information held in working memory during arithmetic problem solving depends on age, context, and problem complexity (for a review of literature pertaining to varying degrees of complexity see DeStafano & LeFevre, 2004), it is unclear whether this conclusions only applies to arithmetic proficiency or if it may be extended to arithmetic fluency, as well.

Inhibition

The term inhibition refers to the executive function responsible for suppressing a previously used or dominant response pattern in order to perform a novel or more difficult response (Miyake et al., 2000). Inhibition may be used to describe a range of characteristics or processes; e.g., to describe a characteristic of one's personality as in social inhibition or to explain a biological action as in a drug inhibitor. However, in the context of executive functions, inhibition is defined as the active suppression of a dominant, i.e., easier to perform/more automatic, response. The active form of inhibition may be divided into two parts: cognitive inhibition and behavioral inhibition (Aron, 2007). Cognitive inhibition, which refers to the active suppression of one's attention to irrelevant stimuli that previously were attended to as well as the active resistance to interference from irrelevant information, is the form of inhibition that is associated with the executive tasks that will be discussed in the coming paragraphs. Behavioral inhibition refers to one's ability to actively control behavior through the suppression of a dominant affective response, such as in the delay of gratification or impulse control. Thus, cognitive inhibition directly contributes to cognitive control while behavioral inhibition contributes to emotion regulation and behavioral control.

The general definition of inhibition is consistent across the range of literature in executive functions; though some may disagree on the specific role of inhibition in the overall framework of executive functions. For example, focus on the "central executive" component of working memory leads to the view that inhibition is a process *of* working memory, rather than simply a process *related* to working memory (e.g., Pennington, 1994). Most likely, this overextension of the definition of working memory to include the processes of inhibition may be attributed to an overgeneralization of the working memory model, a model that encompasses the concept of the "central executive" (Baddeley & Hitch, 1974).

Indeed, studies that have examined the inhibition mechanisms of working memory tend to describe inhibition as a process that is carried out by the central executive. As previously stated, the central executive represents the control system in the working memory model and may be more appropriately defined as an integrated system of multiple cognitive functions, i.e., executive functions, rather than as a defining component of working memory. At the other end of the spectrum lies the view that inhibition – and not working memory – is the underlying component of all forms of cognitive control or executive function (e.g., Aron, 2007). The middle ground between the various theoretical perspectives is represented by the multi-component theories of executive function such as Miyake's three-component model, in which the different executive processes are viewed as three separate but related cognitive functions (Miyake et al., 2000; refer to Figure 2).

Measuring inhibition. Although the general definition of inhibition is consistent across the theoretical literature, researchers often differ in terms of what they consider to be reflections of executive inhibition. Like other components of executive function, these inconsistencies in the literature are due to the lack of a standard methodological approach to assessing cognitive inhibition. Over the years, many measures of inhibition have been developed and utilized; it is not uncommon for a researcher to include multiple measures to tap the underlying construct of cognitive inhibition within a single research design – in fact, this is often preferable. With an ever-increasing number of assessment options and a common desire to include multiple measures, researchers often maintain a degree of consistency across studies by choosing to include at least one out of a small number of traditional measures of inhibition in their research designs.

Some of the more traditional and commonly used measures of executive inhibition are summarized in Table 2 and include such tasks as: the Stroop task (Stroop, 1935), the Stop-Signal task (Logan, 1994), the Erikson Flanker task (Erikson & Erikson, 1974; Erikson & Shultz, 1979), the antisaccade task (Hallett, 1978), and the Go/No-Go task (Donders, 1868/1969). These common measures of inhibition share many methodological and theoretical similarities and often produce statistically similar results. For example, the frequently used Go/No-Go task and the Erikson Flanker task are similar to the Stop-Signal task in that they require participants to ignore irrelevant information and override a dominant response. Miyake et al. (2000) found that the Stop-Signal task, the antisaccade task, and the Stroop task all loaded highly onto a factor for inhibition. Other studies have found evidence to support a similar link between inhibition and the Go/No-Go task (Aron, 2007; Durston et al., 2002) as well as the Erikson Flanker task (Aron, 2007; Bunge et al., 2002; Huizinga et al., 2006; Ridderinkhof & Van der Molen, 1995). It has been suggested that some common executive tasks, such as the Stroop task, the Wisconsin Card Sort Task (WCST), and tasks in task switching require *both* set shifting and inhibition rather than simply one of these processes (e.g., Aron, 2007; Huizinga & van der Molen, 2007). Therefore, although these common measures of executive inhibition undoubtedly tap into mechanisms of inhibitory control, it is important to acknowledge unavoidable issues of task impurity when interpreting measures of executive function. In the current study, the construct of inhibition is defined in terms of performance on the Stop-Signal task and on the inhibition condition of the Shape School task (Espy, 1997) – extended version (Ellefson, Blagrove, Espy, & Chater, 2008).

Age-related changes in inhibition. Regardless of the methodological and/or theoretical discrepancies found within the literature, a common finding in developmental studies of executive function is that processes related to inhibition begin to develop within

Table 2.

Common Inhibition Tasks Grouped by Appropriate Ages

Task	Reference	Task Demand
Child tasks		
Reverse Categorization	Thorndike et al.(1986)	Sort animals into opposite buckets
Shape Stroop	Kochanska, Murray, & Harlan (2000)	Point to small shape embedded in larger shape
Day/Night Stroop	Gerstadt et al. (1994)	Say “night” for sun picture
		Say “day” for moon/stars picture
Grass/Snow Stroop	Carlson & Moses (2001)	Point to white when hearing “grass”
		Point to green when hearing “snow”
Bear/Dragon Stroop	Reed, Pien, & Rothbart (1984)	Follow bear’s instructions
		Do not follow dragon’s instructions
Luria’s Hand Game	Luria, Pribram, & Homskaya (1964)	Make opposite hand gestures
Simon Says	Strommen (1973)	Follow only “Simon’s” instructions
Shape School (original)	Espy (1997)	Name happy faces
		Do not name sad faces
Older child ($\geq$ 6 years) & adult tasks		
Stroop	Stroop (1935)	Say color of ink; do not read word
		Do not read word
Stop-Signal[a,b]	Logan (1984)	Stop responding when tone sounds

Task	Reference	Task Demand
Erikson Flanker [b]	Erikson & Erikson (1974); Erikson & Shultz (1979)	Respond to direction of center arrow Ignore irrelevant arrows
Antisaccade [b]	Hallett (1978)	Ignore distractor stimulus located in one area Respond to target stimulus located in another area
Go/No-Go [b]	Donders (1868/1869)	Frequently respond to target stimuli Inhibit response to rare non-target stimulus
Shape School (extended)[a]	Ellefson et al. (2008)	Respond to happy faces Do not respond to sad faces

[a]Tasks that will be included in the present study. [b]May be used with children ≤ 6 years.

the first year and continue to develop and become more refined over the first two decades of life (Huizinga et al., 2006; Huizinga & van der Molen, 2007; Lehto et al., 2003; Levin et al., 1991). While mechanisms of inhibition are in place by three-four years of age (Espy, 1997), they continue to develop along a protracted course until adolescence or early adulthood. In a sample consisting of 7-year-olds, 11-year-olds, 15-year-olds, and 21-year-olds, Huizinga et al. (2006) found that performance on the Flanker task and the Stop-Signal task reached adult level by 11 years of age. Additionally, they found that 21-year-olds performed significantly better on the Stroop task than any of the younger age groups, indicating that the inhibition processes required in this particular task develop at a slower rate than the processes required in the former two tasks. Their results are consistent with previous studies that have found differential patterns of development associated with different aspects of inhibitory processing (e.g., Dempster, 1992; Nigg, 2000; for a review, see Welsh et al., 2006). Thus, the current evidence indicates that rudimentary inhibitory processes are in place very early in life, experience the most gains between 7 and 11 years, and continue to develop into late adolescence or early adulthood.

Neurocognitive development of inhibition. Like the working memory component of executive function, inhibition has been found to relate to specific patterns of activation in the brain, primarily in the PFC (see Figure 1). From a developmental standpoint, studies have shown that children's brains undergo specific maturational changes that mirror the typical patterns of improvement in performance on tasks requiring executive inhibition. Tamm et al. (2002) examined the performance of children, adolescents, and young adults (ranging in age from 8 to 20 years) on the Go/No-Go task while the participants were undergoing fMRI. This study found that children were more likely to demonstrate patterns of activity in various regions of the PFC while performing the task whereas older subjects demonstrated more confined patterns of activity—specifically in the left inferior frontal gyrus (the orbitofrontal cortex). Concurrently, participants' performance on the Go/No-Go task significantly differed across age groups, indicating that inhibition improved significantly with age (demonstrated by decreased reaction times). Moreover, studies have found that patterns of synaptic development in the PFC mirror patterns of developmental gains in inhibition that are typically observed throughout childhood and into adolescence (Huttenlocher & Dabholkar, 1997; Welsh et al., 2006). Taken together, the results from studies examining this brain-behavior relation have provided consistent evidence favoring the link between cognitive developments in inhibitory processing and age-related changes in brain structure and function.

Inhibition and arithmetic achievement. After working memory, inhibition is the component of executive function that has been the most frequently studied in relation to its role in academic achievement, specifically arithmetic achievement. While several studies have found that inhibition is a significant predictor of arithmetic achievement from preschool to early through late childhood (e.g., Bull & Scerif, 2001; Espy et al., 2004; St. Clair-Thompson & Gathercole, 2006), little is known about the exact nature and stability of this relation beyond childhood. A longitudinal study conducted by Bull and colleagues (2008) sought to carefully examine the dynamic role of the three common executive functions in school achievement from preschool (4-year-olds) to primary school-year three (7- to 8-year-olds). In this study, the researchers assessed children's arithmetic and reading proficiency at three separate time points (preschool, primary school year-one, and primary school year-three) and compared these skills with their working memory skills, inhibition skills, and set shifting skills. They attempted to account for non-executive contributions at each time point by controlling for reading ability when examining predictors of arithmetic performance and controlling for arithmetic ability when examining predictors of reading. Consistent with the results of a similar study conducted by Bull and Scerif (2001), they found that after controlling for reading ability, inhibition and working memory (but not set shifting) accounted for a significant portion of the variance in arithmetic proficiency (measured by accuracy scores on a standardized test of arithmetic) at the end of the participants' final year in preschool. However, at the end of the participants' first year in primary school, inhibition and shifting were not found to predict arithmetic proficiency after reading ability had been controlled. At the end of their third year in primary school, none of the executive measures were significantly related arithmetic. A possible interpretation of these findings is that younger children utilize inhibition when solving simple arithmetic equations because, for them, this activity is still fairly novel. Older children are more familiar with simple equations and may process them automatically, eliminating the need for active inhibition.

Unfortunately, studies examining the relation between academic ability and inhibition have produced conflicting results. St. Clair-Thompson and Gathercole (2006) included multiple measures of working memory, inhibition, and set shifting in their study of 11- to 12-year-old children in an attempt to determine the relation between these common executive functions and performance on school achievement tests in the areas of English, mathematics, and science (again, measures of proficiency). Using a principle component analysis, they were able to isolate two factors that represented inhibition and working memory but their analysis did not identify a component for shifting. However, a similarly designed study conducted with primary school year-four (8- to 9-year-olds) and year-five children (9- to 10-year-olds) used a confirmatory factor analysis to distinguish a factor for working memory and for set shifting but not for inhibition, once they had controlled for naming speed (van der Sluis et al., 2007). The differing results found across studies in this area may be due to age group differences, task (measurement) differences, fundamental developmental differences, or a combination of any of these factors. To date, no two studies have used consistent methodologies to provide a comprehensive analysis of the role of executive inhibition in arithmetic achievement from childhood through adulthood, and no studies have examined the role of inhibition in the context of arithmetic fluency development. Considering that the current evidence indicates that a relation between inhibition and arithmetic proficiency may exist in early and middle childhood, the logical next step for researchers in this area is to refine their methodologies and to expand their developmental perspective with the addition of multiple age groups and multiple measures of arithmetic, i.e., measures of both proficiency and fluency.

Set Shifting

The third component in the three-component model of executive function, set shifting (also referred to as *switching*), entails the active initiation of a new or non-dominant response pattern after the successful suppression of a previous or dominant response pattern (Miyake et al., 2000). In their definition of set shifting, Miyake et al. stressed the importance of defining set shifting as a process that occurs *after* the suppression of an alternate response, illustrating the idea that set shifting is an active and effortful cognitive process rather than a routine or automatic function. To provide a real-world example, in relocating to a new country with a different set of traffic laws, one would utilize inhibition skills to actively inhibit his old driving habits and rely on efficient set shifting skills to properly operate a vehicle under the new set of driving rules.

As previously suggested in relation to measures of inhibition, it is possibly more difficult to isolate the component of set shifting in measurement than the other common executive components such as working memory updating or inhibition. In order to shift from one set of rules or response patterns to another, one also must update relevant information in working memory pertaining to the new pattern, actively refrain from responding to the previous or dominant pattern, and inhibit interference from irrelevant information that would interfere with the successful shifting of mental sets. Consequently, many common measures of set shifting may represent measures of updating and inhibition, and isolating the indicators and influences of fully developed set shifting abilities is often difficult to accomplish.

Measuring set shifting. Miyake and his colleagues (2000) suggested that one method for alleviating some of the problems associated with task impurity would be to include more than one measure of each executive function in a research design and examine the latent common structure among the observed variables. In their study, they chose to measure set shifting using three distinct tasks: the plus-minus task, the number-letter task, and the local-global task. They chose these particular tasks in order to include variety in task requirements and to reduce the likelihood that the shared variance in these tasks would be attributable to similarities in task format or procedures rather than to set shifting ability. A common requirement across all three of these tasks is to shift from one mental set or response pattern to another (e.g., addition or subtraction, letter classification or number classification, and inside shape or outside shape, respectively), as quickly as possible and without making mistakes. They found that the three shifting tasks were more highly correlated with each other than they were with the three inhibition tasks or the three updating tasks. Also, although all three tasks were found to be related significantly to the factor for shifting, the plus-minus task loaded the highest onto this component. These tasks, as well as several other popular measures of set shifting in children and adults are listed in Table 3.

The three set shifting tasks utilized in the Miyake et al. (2000) test battery, though not the three most commonly used measures of set shifting, represent different forms of a particular task paradigm that is routinely used in studies of set shifting. This paradigm, the task switching paradigm (Allport, Styles, Hsieh, 1994; Jersild, 1927; Rogers & Monsell, 1995) along with the Wisconsin Card Sort Task (WCST; Berg, 1948) represent the two most commonly used measures of set shifting in the literature on executive functions in adulthood.

Table 3.

Common Set Shifting Tasks Grouped by Appropriate Ages

Task	Reference	Task Demands
Child tasks		
Standard DCCS	Frye, Zelazo, & Palfai (1995)	Sort cards by color/shape
		Switch to sorting cards by shape/color
Shape School (original)	Espy (1997)	Name objects by their color
		Name objects with hats by their shape
Item-Selection	Jacques & Zelazo (2005)	Pick objects that are similar in a way
		Switch rules and pick objects that are similar in a different way
Older child (≥ 6 years) & adult tasks		
Shape School (extended)[a]	Ellefson et al. (2008)	Name objects by color
		Name objects with hats by shape
Visually Cued Color-Shape (Advanced DCCS)	Zelazo, Craik, & Booth (2004)	Name color if "X" cue (80% of trials)
		Name shape if "Y" cue (20% of trials)
Auditorily Cued Number-Number (Advanced DCCS)	Zelazo et al. (2004)	Name number if shapes in quadrant if male voice (80% of trials)
		Name number in corner of quadrant if female voice (10% of trials)

Task	Reference	Task Demands
WCST	Berg (1948)	Learn sorting rule based on feedback Sort differently when rule changes
Plus-Minus	Jersild (1927); Spector & Biederman (1976)	Perform addition/subtraction Switch between addition & subtraction problems
Number-Letter	Rogers & Monsell (1995)	Number/letter pairs Say “odd” or “even” to *number* if presented in one location Say “vowel” or “consonant” to *letter* if presented in other location
Local-Global	Navon (1977)	Say number of lines in large figure if in blue Say number of lines in smaller embedded picture if in black
Smiling Faces	Rogers & Monsell (1995)	Say gender of figure presented at top of screen Say facial expression of figure presented at bottom of screen
Dots-Triangles	Rogers & Monsell (1995)	Switch between saying number of dots and saying number of triangles

[a]Tasks that will be included in the present study.

Although task switching paradigms have been used with older children (Ellefson, Shapiro, & Chater, 2006; 2010), due to developmental differences in cognitive level, the set shifting abilities of children under the age of six are not properly assessed by way of these popular adult measures. Subsequently, child-friendly executive tasks such as the Dimension Change Card Sort (DCCS; Frye et al., 1995) and the Shape School (Espy, 1997) have provided a viable means to measuring set shifting abilities in young children under six years of age. As the current study involves children older than six, set shifting is measured with a task switching paradigm and with an extended version of the Shape School task (Ellefson et al., 2008).

Age-related changes in set shifting. Performance on developmentally appropriate tasks of executive function indicates that set shifting abilities tend to mature at a slow rate that is similar to rate at which inhibition develops (Huizinga et al., 2006; Diamond, 2002; Welsh et al., 2006). Specifically, several task switching studies have found that shifting performance tends to improve through early and middle childhood until reaching adult level by around 12 years of age (Cepeda et al., 2001; Huizinga & Van der Molen, 2007; Kray et al., 2004), with one study finding that adult level performance may not be reached until early adulthood (Reimers & Maylor, 2005). Upon closer examination of the developmental studies that have been conducted in this area, evidence from studies involving age-appropriate executive tasks indicates that rudimentary set shifting abilities are typically present by five years of age (Frye et al., 1995; Espy, 1997). As tasks increase in difficulty, i.e., the addition of multiple rules and/or multiple dimensions associated with the same stimulus, young children become less capable of successfully shifting between mental sets/responses. Therefore, task/rule complexity seems to be a determining factor in whether or not a child at a given stage of development is able to accurately and efficiently shift between mental sets (e.g., Bunge & Zelazo, 2006).

On a broader scale, general executive tasks that have been postulated to reflect set shifting, such as the Wisconsin Card Sort Task (WCST; Berg, 1948), have also been used to illuminate age-related differences in set shifting. In the WCST, the experimenter instructs a participant to sort each card in a deck of sorting cards according to a common dimension depicted on the cards, i.e., color, shape, or quantity. The experimenter does not tell the participant the correct dimension by which the cards are to be sorted; rather he/she provides the participant with correct or incorrect feedback until the participant has inferred the correct sorting rule. Throughout the course of the task, the participant must infer when the sorting rule has been changed based on the experimenter's feedback. Although set shifting is the executive component that is most highly related to performance on the WCST, evidence indicates that set shifting is not the only executive process responsible for performance on this task (Huizinga et al., 2006; Miyake et al., 2000). Huizinga and van der Molen (2007) assessed age-related changes in the role of executive functions on WCST performance across four age groups (7-year-olds, 11-year-olds, 15-year-olds, and 21-year-olds). They found that set shifting was the strongest predictor of WCST performance at 11 years of age, working memory and set shifting were both strong predictors of performance at 15 years of age, and by 21 years of age, working memory was found to be the strongest predictor of WCST performance. Thus, it seems that set shifting has differential contributing effects on performance on complex executive tasks across development.

As set shifting represents one of the later developing components of executive function, few tasks that are used with adults for the purpose of tapping set shifting are developmentally appropriate for use with young children. Consequently, some researchers have developed their own child-friendly set shifting measures. One such task, similar to the adult WCST, is the Dimension Change Card Sort (DCCS) task that involves sorting cards based on dimension, i.e., the general defining characteristics of the object presented on a card. In this task, children sort cards, one at a time, presented from a stack of cards in one of two locations according to a given rule (Frye et al., 1995). During the first condition, children are instructed to sort cards into their correct piles based on one dimension associated with the object on the card, e.g., shape. In the "shape" game red bunnies and blue bunnies would be sorted into one pile and red trucks and blue trucks would be sorted into another pile. In the second condition, children are told that the rule has changed and they must now sort the cards based on the second dimension, i.e., the object's color. Typically, young children perseverate with the old rule during the second condition – they continue to place the red bunnies with the blue bunnies during the "color game" – until around five years of age (Frye et al., 1995; Jacques, Zelazo, Kirkham, & Semcesen, 1999). The DCCS most likely requires that a child: 1) inhibit the previously relevant dimension (or overcome *attentional inertia* as described by Kirkham et al., 2003), and 2) successfully switch to the new sorting dimension. However, based on the results of the Jacques et al. study involving strategic manipulations of the DCCS, the researchers concluded that young children's perseverative errors more strongly reflect their underdeveloped set shifting abilities than deficits in response inhibition.

Neurocognitive development of set shifting. As with working memory and inhibition, there is compelling evidence demonstrating that patterns of activation in specific areas and regions of the brain are linked to the cognitive processes involved in set shifting. In a recent review, Aron (2008) noted that differential amounts of activation in the right inferior frontal cortex (IFC) directly contribute to the rate at which one is able to stop or slow a response. Furthermore, evidence has shown that activation tends to flow from the IFC to the presupplementary motor area (pre-SMA), and then to the subthalamic nucleus (a region of the basal ganglia). Interestingly, shifting between tasks has been shown to correlate with activity in this network. Individuals with damage to the right IFC have subsequently demonstrated longer reaction times when set shifting and individuals with disrupted pre-SMAs have experienced negative impacts on performance during switch trials versus no impacts on performance during non-switch trials.

Developmental studies that have examined the neurological components of set shifting and the age-related changes that occur across these components have offered additional evidence favoring a link between specific changes in the brain and observed cognitive changes in set shifting abilities. For example, Crone et al. (2006a) examined developmental differences in rule representation and rule switching using a combination of behavioral and neurological data (obtained from fMRI). They compared the brain activity and performance levels of three age groups, 8- to 12-year-olds, 13- to 17-year-olds, and 18- to 25-year-olds, during a standard task-switching task. Consistent with their hypothesis, behavioral data indicated that the two younger age groups performed worse on tasks requiring the use of bivalent rules, i.e., two rules associated with the same stimulus, than tasks involving univalent rules. Neurological data indicated that increased global activation patterns in the youngest age groups during bivalent trials significantly differed from adults who had more specific patterns of activation during these trials. Across all age groups, the researchers found higher rates of activation in the VL-PFC and the superior parietal cortex during bivalent rule use (see Figure 1). However, consistent with behavioral data, children and adolescents significantly differed from adults in amount of VL-PFC activity present while representing bivalent rules. Another age-related difference in activation that emerged in this study related to the pre-SMA. Whereas adolescents and adults tended to activate this region only when switching, children showed patterns of pre-SMA activation during both switching and rule representation. Thus, improvements in set shifting ability typically occurring during late childhood and adolescence significantly relates to increased specialization in the areas of the brain that are responsible for representing and shifting between different mental sets.

Set shifting and arithmetic achievement. Despite the fact that there is increased interest in the relation between school achievement and executive function development, the relation between set shifting and academic performance has garnered the least amount of interest and support in this realm of literature. The studies in this area that have attempted to include a shifting component in their research design have typically involved samples of young children and rarely have attempted to compare the performance of young children with that of older children, adolescents, or young adults. Among those that have involved children between the ages of four and seven, few have found evidence to support a relation between set shifting and arithmetic achievement during this young period of life. For example, Espy et al. (2004) tested preschool children on multiple developmentally appropriate measures of working memory, inhibition, and set shifting and examined the relation between composite scores calculated for each of these components and emergent mathematics skills (as measured by the Woodcock-Johnson, revised Applied Problems subtest – a measure of proficiency). While the evidence from this study indicated that both working memory and inhibition contributed to emergent mathematics skills, the contributions from set shifting were not significant. These results are consistent with the results from Bull et al. (2008) who were unable to find a significant relation between set shifting and arithmetic proficiency after controlling for reading ability in 4-5-year-olds, 5-6-year-olds, and 7-8-year-olds. Although Espy et al. suggested these results may be due to the later development of set shifting or the minimal level of complexity associated with the simple arithmetic problems that are appropriate for younger children, evidence from studies involving older children have not consistently supported this claim.

The majority of the studies in this area that have involved school-age children have not found a consistent association between set shifting and many areas of mathematics achievement. Two similar studies involving 11- to 12-year-olds found that set shifting did not relate to proficiency in solving arithmetic word problems (Lee et al., 2009) or to proficiency in a school-based measure of mathematics achievement (St. Clair-Thompson & Gathercole, 2006). However, a study conducted with 6- to 8-year-old children found that set shifting abilities, as measured by performance on the WCST, did provide a significant contribution to arithmetic proficiency even after controlling for reading ability and intelligence level (Bull & Scerif, 2001). However, the findings from this study may be due to developmental differences and/or differences in arithmetic assessment. In other words, preschool children are still developing the skills needed to perform arithmetic procedures. Through time and experience, these children move beyond simple competence and into more complex processing. Thus, 6-to 7-year-olds may require shifting skills when they need to move beyond the level at which they can simply perform addition and subtraction and on to the level at which they are able to efficiently alternate between using addition and subtraction. Once children become proficient in performing basic mathematical operations (by around 11 years of age), they may start to switch between these operations rather automatically, i.e., fluently. In any case, considering the lack of evidence, one cannot definitively conclude that set shifting does or does not contribute to arithmetic proficiency and fluency without first comparing this relation across multiple age groups using consistent methods and including measures of both proficiency and fluency.

Summary

The reports that have been summarized throughout this review provide evidence of a link between executive functions and arithmetic proficiency in the preschool years (Bull & Johnston, 1997; Espy et al., 2004), early and middle childhood (Bull et al., 2008; Bull & Scerif, 2001; Gathercole & Pickering, 2000; Gathercole et al., 2004; van der Sluis et al., 2007), and during the middle school years (Lee et al., 2009; St. Clair-Thompson & Gathercole, 2006). Testing a more developmentally diverse sample on measures of both proficiency and fluency would provide evidence of the relations between age, executive functions, and achievement across multiple age groups rather than within a particular age group and within one arithmetical context. So far, evidence has supported the relation between arithmetic achievement and working memory in preschoolers (Bull et al., 2008; Espy et al., 2004), young children (Bull et al., 2008; Gathercole et al., 2004b), and older children (Gathercole et al., 2004; Gathercole & Pickering, 2000; St. Clair-Thompson & Gathercole, 2006). There is some evidence to indicate a relation between inhibition and arithmetic achievement exists in preschool children (Bull et al., 2008; Espy et al., 2004), primary school children (Bull & Scerif, 2001), and older children (St. Clair-Thompson & Gathercole, 2006). However, it is worth noting that Bull et al. (2008) found that this relation did not persist from the beginning of primary school to the end of the children's first year in primary school nor did it persist into their third year of primary school, a finding that is consistent with van der Sluis et al. (2007) but inconsistent with St. Clair-Thompson and Gathercole (2006). Although the results are mixed regarding the role of set shifting in arithmetic achievement – with several studies finding evidence against this association (Bull et al., 2008; Espy et al., 2005; Lee et al., 2009; St. Clair-Thompson & Gathercole, 2006) and others finding evidence in favor of it (Bull & Scerif, 2001; van der Sluis et al., 2007) – lack of consistency in age group comparisons and measures used may account for these differences. In an attempt to partially alleviate such discrepancies and to gather more information regarding the relations among executive functions, arithmetic proficiency, *and* arithmetic fluency across a wider span of ages, the current study focused on the role of each executive function component in performance on a measure of arithmetic proficiency and a measure of arithmetic fluency in children ages 6-7 years, 9-10 years, adolescents ages 12-13 years, and adults ages 18 years and older.

Purpose of the Study

This study investigated the predictive value of age, working memory updating, inhibition, and set shifting to arithmetic, in general, and to arithmetic proficiency and fluency, specifically, while controlling for verbal and non-verbal general cognitive ability. As one would expect to find general age-related differences in the academic domain of arithmetic (Butterworth, 2005; Prather & Alibali, 2009), it was expected that the older participants in this sample would demonstrate better arithmetic performance than the younger participants. In addition, based on previous findings, it was expected that the results would support age-related differences in executive function skills in the areas of working memory updating (Huizinga et al., 2006; Luciana et al., 2005), inhibition (Huizinga et al., 2006), and set shifting (Cepeda et al., 2001; Kray et al., 2004) such that the older participants would demonstrate greater skills in these areas than the younger participants. A positive linear relation between executive function skills and arithmetic performance also was expected, regardless of age (Bull et al., 2008; Bull & Scerif, 2001; St Clair-Thompson & Gathercole, 2006), for both proficiency and fluency. Thus, the study examined the partially mediating role of executive functions on the relation between age and arithmetic.

Although this study hypothesized that age would affect arithmetic performance, due to lack of statistical power, this study did not explore the possible moderating effects of age on the relation between executive functions and arithmetic; rather, the mediating effects of differences in executive function skills on the relation between age and arithmetic was explored. Baron and Kenny (1986) distinguished between moderating and mediating effects by defining a moderator as a variable that's state affects the direction and/or strength of the relation between two other variables and a mediator as a variable that accounts for how or why the relation between two other variables exists. In a typical moderating situation, both the predictor variable and the moderator variable directly affect the outcome variable, i.e., Predictor → Outcome and Moderator → Outcome, and the interaction between the predictor and the moderator also influences the outcome, i.e., Predictor X Moderator → Outcome. In the current study, moderation could have been tested using multi-sample path analysis (had the sample size been larger), which would have produced separate path diagrams for each age group. In a mediating situation, three conditions must hold: (1) the predictor variable must be related to the mediator variable such that changes in the predictor account for changes in the mediator, (2) the mediator variable also must be related to the outcome variable such that changes in the mediator account for changes in the outcome, and (3) the relation between the predictor variable and the outcome variable through the mediator variable must account for all or part of the direct relation between the predictor and the outcome, i.e., the inclusion of the meditational paths reduce or eliminate the direct effects of the predictor on the outcome. As previously mentioned, one of the purposes of the current study was to examine the partially mediating effects of differences in executive functions on the association between age-related differences and differences in arithmetic performance, as it was expected that age-related differences in executive function (due to documented developmental differences in executive functions) would account for a portion of age-related variance in performance.

In addition, this study examined the contributions of executive functions to overall arithmetic performance, and to arithmetic proficiency and fluency, after controlling for differences in age and general cognitive ability. The presumed distinctiveness of intelligence or general cognitive ability and executive functions was expected to allow for the relation between executive functions and arithmetic performance to remain significant after accounting for individual differences in both verbal and non-verbal general cognitive ability (Friedman et al., 2006). In examining the cognitive and developmental predictors of arithmetic performance across a wide range of ages, and by including measures of both proficiency and fluency, the results from this study contribute to the growing body of knowledge on the cognitive components that contribute to the development of multiple forms of arithmetic skill across grade levels, and may inform future efforts that aim to improve such skills.

Hypotheses

1. A significant positive linear relation exists between arithmetic performance and: (a) age,

(b) general cognitive ability, (c) working memory updating, (d) inhibition, and (e) set shifting. In other words, older participants as well as participants with higher scores on the measures of general cognitive ability, working memory updating, inhibition, and set shifting demonstrate better arithmetic skills (see Figure 6).

2. A significant positive linear relation exists between age and: (a) general cognitive ability, (b) working memory updating, (c) inhibition, and (d) set shifting. The older participants demonstrate better performance on measures of general cognitive ability and executive functions.
3. Working memory updating, inhibition, and set shifting account for a significant portion of variance in arithmetic scores above and beyond the influence of (a) age and (b) general cognitive ability.
4. Together, age, general cognitive ability, and executive functions simultaneously explain a significant portion of variance in arithmetic performance, such that the inclusion of each component provides a significant contribution to the overall fit of the prediction model.

Method

Participants

The participants included in this project were part of an existing data set (Ellefson et al., 2010). A total of 148 participants participated: 36 primary school Year-2 children (17 males, 19 females) and 44 primary school Year-5 children (26 males, 18 females) from two primary schools in England; 36 secondary school Year-8 children (17 males, 19 females) from one secondary school in England; and 32 adults (10 males, 22 females) from a neighboring university in England (see Table 4). From the full data set, two Year-2 children with special needs and one university student with very limited English proficiency

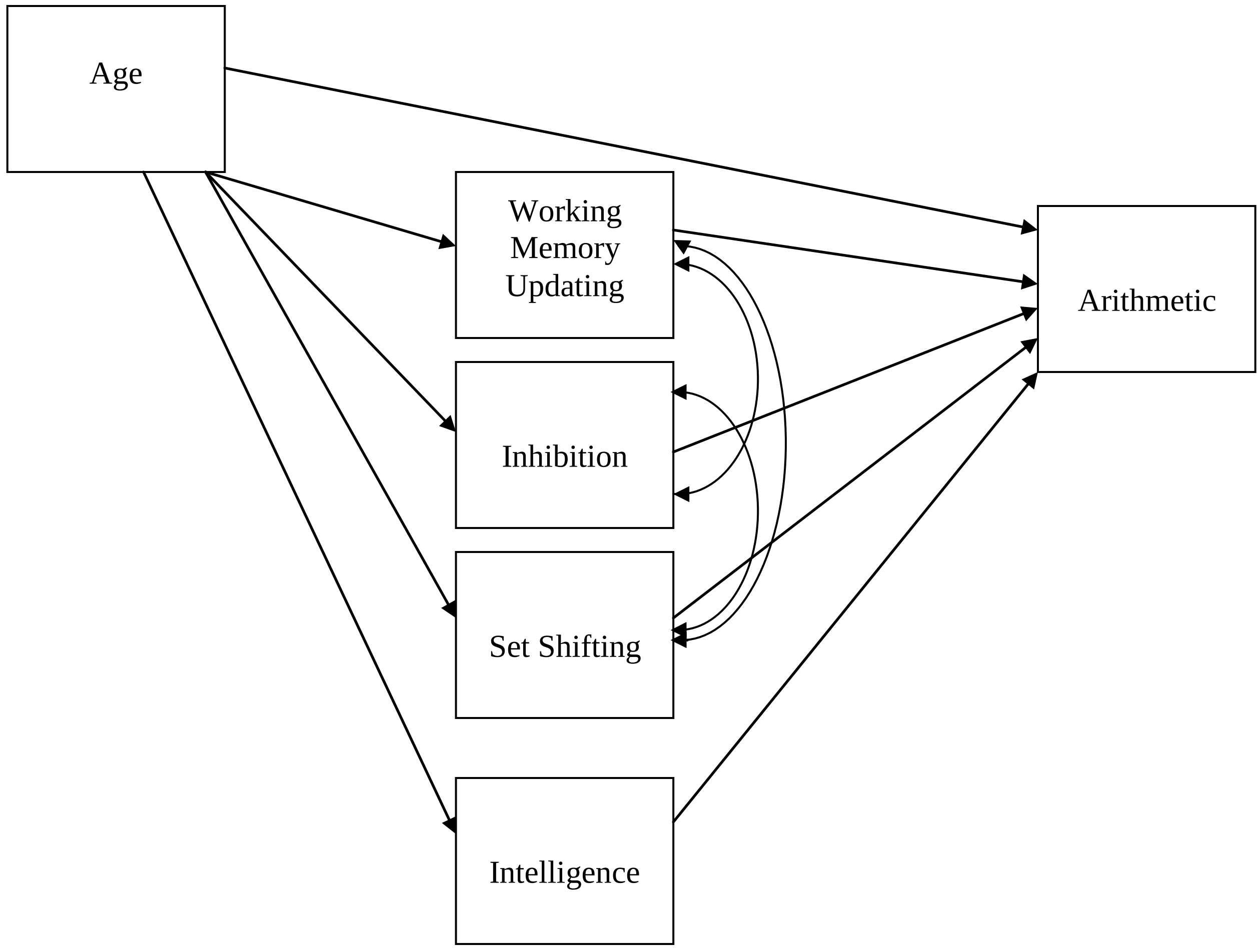


Figure 6. Hypothesized path model connecting all IVs with the DV, showing the correlations among the three executive components, and displaying the partially mediating role of executive functions and intelligence in the relation between age and arithmetic performance.

Table 4.

Participant Ages (in Years) by Age Group and Gender (n = 148)

	n (Males)	*M*	*SD*	Range
Age Group				
Year 2	36 (17)	7.08	.33	6.62 – 7.61
Year 5	44 (26)	10.03	.30	9.56 – 10.53
Year 8	36 (17)	13.04	.26	12.59 – 13.51
University	32 (10)	24.16	5.15	18.56 – 39.54

participated in the study but were unable to complete all of the tasks due to certain task demands. These three participants were not included in the current sample description or in the subsequent analyses. School and parental consent and participant assent were obtained for all participants under 18 years of age and consent was obtained from participants who were 18 years of age or older. Ethnicity and social economic status were not collected from participants; however, the sample was representative of the region where the data were collected.

Measures

All participants completed six executive function tasks that were chosen to represent working memory updating, inhibition, or set shifting, two standardized measures of general cognitive ability, one measure of arithmetic proficiency, and one measure of arithmetic fluency as part of participating in a separate study on age-related differences in task switching (Ellefson et al., 2010, see Table 5 for a summary). This section describes how each of those tasks was administered in that project and the variables from those tasks that will be used in this study.

Table 5.

Cognitive Constructs and Independent Variables (IVs) From Each Measure

Construct	Measures	IVs (ABBREVIATION)
Working Memory Updating	1. Recall of Digits – Forward	1. Raw Score (WM-DIGITS)
	2. Tic-Tac-Toe	2. Efficiency (WM-TTT)
Inhibition	1. Shape School – Inhibition	1. Efficiency (INH-SS-INH)
	2. Stop-Signal	2. Stop-Signal RT (INH-SSRT)
Set Shifting	1. Shape School – Switch	5. Efficiency (SS-SS-SW)
	2. Figure Matching	6. Efficiency (SS-FIGURES)
Intelligence	1. BPVS	1. Raw Score (GCA-BPVS)
	2. Raven's	2. Raw Score (GCA-RAVENS)
Arithmetic	1. WRAT-III	1. Raw Score (MATH-WRAT)
	2. Task Switch - Math	2. RT to accurate trials (MATH-TSMATH)

Note. The abbreviated titles assigned to each IV are important as they will be utilized in describing the results.

Executive functions. ***Working memory updating.*** Two tasks were used to assess working memory updating, the Recall of Digits subtest-Forward of the British Abilities Test - 2nd edition (BAS-II; Elliot, 1996) and the Tic-Tac-Toe task (Huizinga et al., 2006; adapted from Milner, 1971).

Digit span tasks test a person's ability to recall and repeat increasingly long stings of digits in the correct serial order. Several of the most common measures of cognitive ability include a digit span subtest to measure working memory ability; for example, the Wechsler Adult Intelligence Scale-fourth edition (WAIS-IV; Wechsler, 2008) and the Wechsler Intelligence Scale for Children-fourth edition (WISC-IV; Wechsler, 2003) both include digit span subtests. There are two forms of digit span tasks – forward digit span and backward digit span – and both fall into the category of simple span tasks, i.e., they involve single-digit numbers rather than complex lists of items (as opposed to complex span tasks; see Daneman & Carpenter, 1980 and Vock & Holling, 2008). The current study included a forward digit span task, the Recall of Digits: Forward subtest of the BAS-II. The BAS-II is a standardized measure of preschool (from 2 years, 6 months) and school-aged (from 17 years, 11 months) children's overall and specific cognitive abilities and includes six core cognitive subtests, three achievement tests, and five cognitive diagnostic subtests (including Recall of Digits: Forward). The Differential Abilities Scale (DAS; Elliott, 1990) was created as a North American equivalent to the BAS (Elliot, 1983); the Recall of Digits: Forward task on both tests is identical.

Using the standard procedure of the Recall of Digits: Forward subtest of the BAS-II, participants heard a string of digits at two digits per second and repeated the digits (orally) at the end of each string. As the test progressed, the digit strings became increasingly longer until the participant reached ceiling, i.e., more than one mistake or pass in a block. A high final score on this measure indicates that one has a high ability to store, update, and recall simple information. The average split-half reliability score that has been reported for the overall measure (the BAS-II) is .85 – based on a sample ranging from 6 years of age to 17 years of age – and Elliot (1996) reported a high correlation between the BAS-II Recall of Digits-Forward subtest and the Digit Span subtest of the Wechsler Intelligence Scale for Children, third edition (WISC-III; Wechsler, 1991), r = .68. The variable that was taken from this measure was participant's raw score.

In the Tic-Tac-Toe task (Huizinga et al., 2006; adapted from Milner, 1971), participants memorized a pattern of three Xs and Os, i.e., low working memory load, or four Xs and Os, i.e., high working memory load, presented in a 3x3 grid; the number of letters per given pattern, i.e., three or four, varied across trials (see Appendix A). Once participants were comfortable with the pattern, they initiated the recognition phase of the task by pressing the space bar on the computer's keyboard. During the recognition phase, a series of Xs and Os appeared one at a time for 600 milliseconds (ms) each in one of the nine spaces within the 3x3 grid. The low memory load trials include four to seven of these presentations and the high memory load trials include four to nine presentations. The participants pressed a button when all of the components present in the initial pattern had been presented on the screen. Participants completed two practice trials: 15 low memory load trials, and 15 high memory load trials (the order of the low and high memory load trials was counterbalanced across participants). Reliability and validity information for this task is unavailable as it is an experimental measure. Huizinga et al. (2006) reported that accuracy on this task loaded highly onto the same factor (working memory) as accuracy on a mental counters task (Larson et al., 1988), another task of working memory. Accuracy and reaction times (RTs) were collected for each participant per each trial, and the variable taken from this task was accuracy divided by RT (efficiency).

Inhibition. The inhibition tasks used in this study included the Stop-Signal task

(Logan, 1994), and the Shape School (Espy, 1997) – Extended version (Ellefson et al., 2008).

The computer-based Stop-Signal task (modified for children from Logan, 1994) used here was a visual choice reaction time task, as participants were visually presented with opposing stimuli and had to choose the appropriate response for each stimulus presented (van Boxtel, van der Molen, Jennings, Brunia, 2001). While participants were performing the visual choice task, a signal was emitted intermittently following a stimulus presentation. Participants had to refrain from responding when trials were accompanied by the signal. During this task, the delay period between the presentation of a stimulus and the emission of the stop signal tone was continuously adjusted depending on the participant's performance. When the delay period increased, it was more difficult to correctly inhibit a response (van Boxtel et al., 2001). Thus, if a participant was performing well, the delay period continued to increase until a mistake was made, i.e., the participant incorrectly responded to an inhibit trial, and decreased in response to poor performance. As the speed at which an individual provided a response as well as the average delay period between the stimulus presentation and the stop signal were both important factors of performance on this task, the main dependent variable taken from this study, stop signal reaction time, incorporates both reaction time and average delay length. The Stop-Signal task provides a good measure of inhibition as it requires the active suppression of a preponent response (Logan, 1994).

The Stop-Signal task, administed by the E-Prime® stimulus presentation program, (Schneider, Eschman, & Zuccolotto, 2002) required participants to indicate the direction of the arrow (either left- or right-facing) that appeared on the computer screen by pressing either the m-key (to respond to right-facing arrows) or the z-key (to respond to left-facing arrows) on the computer's keyboard while inhibiting their responses when the arrows were accompanied by an audible tone, i.e., stop-signals (the trials that were not accompanied by a tone are referred to as “go” trials and the trials with a tone are referred to as “stop” trials). Both the m-key and the z-key were pre-marked with yellow stickers. Trial stimuli are provided in Appendix B. The stop signal was a computer-emitted 1,000-hertz tone that was played for up to two seconds (the tone stopped when the participant responded). There were four practice trials: two go trials and two stop trials. The experimental trials were presented in three 32-item blocks, for a total of 96 test trials. Within each block, there were 24 go trials (75 percent of trials) and 8 stop trials (25 percent of trials), presented in random order. The stop signal occurred at random during 24 of the experimental trials, i.e., the inhibit trials (during one-fourth of the total trials). Participants were asked to respond to each trial as quickly and as accurately as possible. The E-Prime® program recorded RT and accuracy data for each trial. Miyake et al. (2000) reported a high split-half correlation reliability estimate for this measure (r = .92) with adult participants, and found that the stop-signal variable loaded onto the same factor (factor loading = .33) as two other common tasks of inhibition, an antisaccade task (factor loading = .57) and a Stroop task (factor loading = .40).

The variable that was taken from the Stop-Signal task, Stop Signal Reaction Time (SSRT), was calculated for participants by subtracting their critical Stop Signal Delay (SSD) by their mean go trial RT based on Equation 1 (used by Ray Li, Huang, Constable, and Sinha, 2006).

$$\text{Mean } RT_{\text{Go Trials}} \quad (1)$$

The SSD represents the time interval between the presentation of a go signal and the presentation of the stop signal. The SSD varied across trials, according to participant performance, using the staircase procedure. With the staircase procedure, the SSD started at

350 ms and either increased or decreased by 50 ms on each subsequent trial depending on whether the participant correctly inhibited a response – resulting in an increase – or failed to inhibit a response – resulting in a decrease. The critical SSDs that were used to compute each participant's SSRT represent the SSD at which the participant was able to inhibit a response approximately half of the time (Levitt, 1971). Participants with larger SSRTs were able to tolerate longer delays between go signals and stop signals while still correctly inhibiting a response, and required less time to respond in go trials.

The second measure used to assess inhibition was an extended version of the Shape School (Espy, 1997; Ellefson et al., 2008), a task that was originally designed for young children (see previous sections, *Measuring Inhibition* and *Measuring Set Shifting*). The extended version followed the same general format as the original version but was designed for use with older children and adults and contained 48 trials per condition, compared to 15 in the original version (Ellefson et al., 2008). As previously described, the Shape School is a paper-based task that uses a storybook format in order to measure different aspects of executive function. The task includes four conditions, Control, Inhibition, Switch, and Both. All participants in this study completed the four conditions in the same order (Control-Inhibition-Switch-Both). At the beginning of the task, participants were introduced to the children of the Shape School – either red or blue, circle or square, cartoon figures with neutral facial expressions, two arms, and two legs – and told that the children's names correspond with their color. Each condition contained eight rows of figures with six figures on each row.

The original version of the Shape School task (Espy, 1997) was designed to measure inhibition and set shifting in young children. In this storybook task designed for children as young as four years of age, the experimenter introduces the Shape School and instructs the child to name the school children as they progressed through their daily activities. The "children" in the original version of the Shape School are either circles or squares, red, blue, or yellow. There are four conditions in the measure, each consisting of three separate lines of five distinct stimuli (trials). In the first condition, participants name the school children in Mr. Circle's and Ms. Square's classes one-by-one, as quickly and accurately as possible, according to their color. This condition is meant to familiarize participants with the task and assess their baseline performance level. The second condition, the inhibition condition, includes trials that alternate between non-inhibit trials and inhibit trials. The participants are instructed to name the "children" who have happy faces – still based on their color – and refrain from naming the children who have sad faces, i.e., the inhibit trials. Thus, participants need to actively stop themselves from performing a practiced and continuous response in order to perform the task correctly. The next condition will be described in a later section pertaining to measures of set shifting. Although the demands of this task would be challenging for individuals of all ages, the restricted number of trials in each condition of the original version does not capture the range of abilities of individuals older than around six years of age. Thus, Ellefson et al. (2008) designed the extended version to be used with older children and adults. The only differences between the extended version and the original version are in the number of trials and the color of the stimuli used (i.e., the extended version uses only red and blue). Due to the age range of the sample in the current study, inhibition was measured by performance on the extended version of the Shape School task.

The Inhibition block of the extended version of the Shape School was the second condition in the task and followed from the Control block. Trials corresponding with happy faces were labeled “go trials” and trials containing sad faces were “inhibit trials”. This block contained 24 go trials (12 red and 12 blue) and 24 inhibit trials (12 red and 12 blue), in randomized order across the rows. Scoring was based on the first response provided for each trial. Handheld stopwatches were used to assess participants’ time to complete all 48 trials, estimated to the nearest second, per each block. The extended version of the Shape School is still experimental, so reliability and validity data are not yet unavailable. However, based on a sample of 219 children ages three to six years, Espy et al. (2006) reported high internal validity for the Inhibition block of the original Shape School, $\alpha = .71$, and moderate predictive validity of latency scores on the Inhibition block of the Shape School to raw scores on a standardized measure of inhibition, the Visual Attention task (Korkman, Kirk, & Kemp, 1998), $r = -.22$, $p < .01$ (Espy, Bull, Martin, & Stroup, 2006). Both forms of data that were collected from each participant on the Inhibition block, i.e., RT and accuracy, were used as a reflection of inhibition skills in the analyses by dividing accuracy by RT (efficiency).

Set shifting. Set shifting was assessed using the Shape School – extended version (Espy, 1997; Ellefson et al., 2008) and a figure matching task-switching paradigm (Ellefson et al., 2006).

After the Control block and the Inhibition blocks in the Shape School (Espy, 1997), participants completed a third condition: the Switch block. The premise of the Switch block was similar to the premise of the first two conditions (see previous section in Methods, *Inhibition*) but involved a different naming rule. In the this block, participants were told to name the “children” in Ms. Hat’s class, the children with hats on, by their shape (the switch trials) and to continue to name the children from Ms. Square’s class by their color (the repeat trials). The trials in this condition alternated at random between shape (with hats) trials and color (without hats) trials and between repeat (shape-shape or color-color) and switch (shape-color or color-shape) trials. Thus, participants had to shift between mental sets intermittently throughout this condition in order to perform well. After the Switch block, the fourth condition involved a combination of inhibit, switch, and repeat trials, i.e., it included a combination of happy and sad children with and without hats. This condition is often very difficult for children and adults alike to perform; because of its complexity, it was not included in the analyses of the current study.

The Shape School task is an excellent example of a task that was created specifically to measure both inhibition and set shifting in younger populations. Since the creation of the task in 1997, an increasing number of developmental studies examining executive functions have included one or more measures from the Shape School in their research design (e.g., Bull et al., 2008; Senn, Espy, & Kaufmann, 2004) and a study examining the validity of the measure has found support for its efficacy (Espy et al., 2006). As previously described, the extended version of this task (Ellefson et al., 2008) is more suitable for use with older children and adults and was used as a measure of set shifting in the current study.

In the current study, participants named each figure with a hat on by its shape (a total of 24 trials, 12 circles and 12 squares) and each figure without a hat by its color (a total of 24 trials, 12 red and 12 blue) as quickly and as accurately as possible. Half of the figures had hats and the other half did not. Espy et al. (2006) reported high internal validity for the Switch block of the original Shape School (Espy, 1997), $\alpha = .80$, and moderate predictive validity of latency scores on the Switch block of the Shape School to raw scores on the Visual Attention task (Korkman et al., 1998), $r = -.21$, $p < .01$. The scoring procedure and materials that were used on this block were the same as those used in the Inhibition block, and accuracy on the Switch block divided by RT (efficiency) was used as the variable reflecting set shifting skills.

The task switching paradigm typically involves switching between *pure* task blocks, i.e., blocks that require only one stimulus-response set ("AAAA"), and *alternating* blocks, i.e., blocks that require alternating between pure and switch trials ("AABBAA"), or *mixed* blocks, i.e., blocks that require participants to shift response sets with every trial ("ABABA"). Across the adult (Meiran, 1996; Meiran, Gotler, & Perlman, 2001; Rogers & Monsell, 1995) and developmental task switching literature (Cepeda, Kramer, & Gonzalez de Sather, 2001; Dibbets & Jolles, 2006; Kray, Eber, & Lindenberger, 2004), researchers have suggested that efficient switching, defined as the ability to rapidly and accurately alternate between responses on a trial-by-trial (A to B) and block by block (block A to block B) basis, depends on the active updating and initiation of new response sets. Although there are numerous studies related to the factors that contribute to task switching in adult populations and relatively few developmental studies by comparison, interest in the developmental components that contribute to age-related differences in task switching has been growing steadily over the past decade and is likely to generate new insight on human cognitive development within the coming years.

One of the measures of set shifting that was used in the current study was the Figure Matching task-switching paradigm (Ellefson et al., 2006; 2010), a recently developed measure that reflects the common purpose of general task switching paradigms. In this measure, participants were presented with one of two shapes (either a triangle or a circle) that appeared in one of two colors (either red or blue) in the center of a white computer screen, and were instructed to match each center shape with one of two smaller shapes on the bottom of the screen based on either color or shape, as indicated by a rule cue. In support of previous studies (e.g., Jersild, 1927; Rogers & Monsell, 1995), Ellefson and colleagues (2006; 2010) found that when participants were engaging in set shifting, their reaction times were longer than when they were performing repeat responses, indicating that the active shifting of mental sets requires the expenditure of cognitive resources not typically required during automatic processing. Task switching researchers have suggested that "switch costs," or the carry-over cognitive effects of completing a switch trial as reflected by longer reaction times on the subsequent trial, reflect the collective activation of inhibition and set shifting processes and not simply set shifting, alone (e.g., Mayr, 2002; Meiran, 1996).

In the computer-based figure matching exercise, participants responded to the center stimulus by pressing either the right key (the m-key on the keyboard marked with a yellow dot sticker) or left key (the z-key on the keyboard marked with a yellow dot sticker), on the computer's keyboard, depending on the trial cue. The correct matching task was provided by the rule prompt (or "cue") that appeared at the top of each stimulus slide (see Appendix C). There were 8 practice trials (not included in data analysis) and 100 test trials. The test trials were administered over a series of four 25-item blocks that occurred in random order: (1) a pure color block, (2) an alternating runs block, (3) a pure shape block, and (4) a mixed block. During the pure blocks, trials included the same task (selecting matching shape or color). In the alternating runs block, trials alternated between the two tasks in a predictable order, i.e., shape-shape-color-color-etc. In the mixed block, participants switched tasks after every trial, i.e., shape-color-shape-color-etc. The figure matching stimuli included two different shapes (triangles and circles) that were presented in one of two colors (red or blue) on a white background (see Appendix C). The E-prime® stimulus presentation program (Schneider et al., 2002) administered task instructions, presented stimuli, and recorded participants' accuracy and RTs (ms) for each test trial. As the figure matching task is in its experimental stages, reliability or validity data are not available. However, this standard paradigm is part of a large experimental literature (see Monsell, 2003 for a review). Participants' RTs for accurate trials and accuracy on the alternating runs block were used to form the efficiency variable taken from this task (accuracy over RT).

General cognitive ability. To account for a potential source of individual variation, two measures of general cognitive ability were administered to all participants. Non-verbal ability was measured through a pattern recognition task, Raven's Standard Progressive Matrices (Raven, Raven, & Court, 1993). Verbal ability was measured using the British Picture Vocabulary Scale-Second Edition (BPVS-II; Dunn, Dunn, Whetton & Burley, 1997), the British equivalent of the Peabody Picture Vocabulary Test (PPVT; Dunn & Dunn, 2007).

Raven's Standard Progressive Matrices (Raven's; Raven et al., 1993, updated 2007) is a standardized measuring of reasoning abilities for individuals aged six years and older. Following the test instructions provided in the manual, participants completed a total of 60 trials, grouped units of 12 across five sets (from Set A to Set E). With each trial, participants were shown a picture of an object with a missing piece and told to choose the piece that completed the object among the six or eight items that were pictured below the center item. An experimenter recorded the responses of Year-2 and Year-5 children, and participants in Year-8 or older recorded their own responses. This task was administered individually for Year-2, Year-5, and university participants and administered in groups for Year-8 participants. The internal consistency reliability estimate for Raven's, based on the standardization sample of 793 individuals, has been reported to be high (.88; Raven et al., 1993, updated 2007). In addition, the convergent validity between Raven's and full-scale IQ scores on the Weschler Adult Intelligence Scale-Revised (WAIS-R; Weschler, 1981) has been reported to be in the range of .74 to .84 (O'Leary, Rusch, & Guastello, 1991). The variable used from this measure was participant's raw score – reflecting the total number of trials to which the participant responded correctly.

The BPVS-II (Dunn et al., 1997) is a standardized measure of verbal ability with a similar multiple-choice format; however, unlike Raven's, it is intended to measure receptive vocabulary. In each trial of this task, the experimenter verbally states a vocabulary word and the participant selects one of four pictures that best represents that word. The test contains a total of 14 sets (with 12 items each) that range in degree of difficulty and progress from least difficult to most difficult. Participants began with the set number that corresponded with their age and then continued on to the next set. If a participant made more than one error in the first set administered, the experimenter administered the previous set(s) until the participant completed a full set with no more than one error (until the participant had established a basal set). Participants progressed through all 12 items in each set until they made eight or more errors in a single set (until they reached ceiling). The manual for the BPVS-II reports a high split-half reliability estimate for this measure, $r = .86$, based on the normative sample of individuals from age 3 to age 15. In addition, Elliot (1983) reported that the original version of the BPVS was moderately correlated with the BAS Word Reading Test, $r = .51$, and to general intelligence as measured by the BAS, $r = .60$. The variable used from this measure was participant's raw score – or total number of correct responses.

Arithmetic. An arithmetic switching task (Ellefson et al., 2006; 2010) was used to measure arithmetic fluency, and the arithmetic portion of the Wide Range Achievement Test-

Third Edition (WRAT-III; Wilkinson, 1993) was used to measure arithmetic proficiency.

The arithmetic switching task (task switch math; abbreviated TSMath) was consistent with the figure matching task that was previously described with the exception of the stimuli involved (Ellefson et al., 2006; 2010). Here, instead of responding to the shape or color of the center stimulus, each trial included simple arithmetic problems (single-digit addition and single-digit subtraction), also located on the center of the screen (see Appendix D). This task contained a total of eight practice trials and 100 test trials, divided into four 25-item blocks. The two pure blocks contained either all addition trials or all subtraction trials. The alternating runs block alternated among two-trials runs of addition and of subtraction, e.g., + + - - + + - -. The trials in the mixed block alternated between addition trials and switch trials and did not include repeat trials, e.g., + - + - + - + -. The E-prime® stimulus presentation program (Schneider et al., 2002) administered task instructions, presented stimuli, and recorded participants' accuracy and RTs (ms) for each test trial. As with the figure matching task, the TSMath task is an experimental task using an established experimental paradigm (Monsell, 2003), thus reliability and validity data are not available. The fluency variable used from this task was the average reaction time for all accurate trials.

The standardized version of the WRAT-III (Wilkinson, 1993) is a brief achievement test comprised of three subtests: reading, spelling, and arithmetic designed to be completed in less than 30 minutes. The arithmetic switching task contained only single-digit addition and subtraction problems, but the WRAT-III contains an untimed oral section including more complex number computations and orally-dictated word problems, and a timed written section in which problems increase in difficulty (from simple to complex computations). The arithmetic subtest was administered according to the instructions provided in the test manual.

Year-2 children completed both the oral and the written components while participants in Year-5 or higher completed only the written component (unless, however, a participant provided five or more incorrect responses on the written component; then that participant completed the oral component as well). The written component consisted of 40 computational problems that participants were given 15 minutes to complete. Reliability for this measure is based on the normative sample that included individuals ranging from 5 to 74 years of age, Cronbach's $\alpha = .85$ (Wilkinson, 1993). In terms of the validity of the measure, Wilkinson (1993) reported a high correlation between performance on the arithmetic subtest of the WRAT and scores on the arithmetic portion of the more extended WISC-III (Weschler, 1991), $r = .67$. The proficiency variable used from this measure was participant's raw score (maximum total score = 55).

Procedure

All participants completed the full battery of tasks over multiple sessions in a quiet room of their school as part of a separate study (Ellefson et al., 2010). Four of the tasks were administered by computer (Tic-Tac-Toe, Stop-Signal, Figure Matching, and Arithmetic Matching), two were paper-based (WRAT-III and Raven's), and three were administered orally by an experimenter (Digit Span, Shape School, and BPVS-II). The E-Prime® program (Schneider et al., 2002) was used to administer task instructions and present the stimuli for all of the computerized tasks. Children completed the computerized tasks while seated in front of a laptop computer and university students completed these tasks while seated in front of a desktop computer. Upon completion of each testing session, children were rewarded with stickers and college students received £15 (British pound sterling, roughly equivalent to $25). The entire task battery took approximately 1.5 hours to complete and the order of the task series was counterbalanced across participants.

Results

Description of Data, Screenings and Transformations

Prior to conducting any analyses, all variables that were of interest from each of the tasks were screened for violations of the assumptions of multivariate regression and path analysis. There were 10 measured variables: Digit Span (*WM-DIGITS)*, Tic-Tac-Toe (*WM-TTT)*, Shape School-Inhibition Block (*INH-SS-INH*) Stop-Signal (*INH-SSRT*) Shape School-Switch Block (*SS-SS-SW*) figure matching (*SS-FIGURES*) BPVS-II (*GCA-BPVS*) Raven's Progressive Matrices (*GCA-RAVENS*) WRAT-III (*MATH-WRAT*) and task switch-math (*MATH-TSMATH)*; also, one variable was included for age (*AGE)*. Measured variables were standardized and averaged to form five composite variables: WM-DIGITS and WM-TTT (*WMU*), INH-SS-INH and INH-SSRT (*INH*), SS-SS-SW and SS-FIGURES (*SHIFT*), GCA-RAVENS and GCA-BPVS (*GCA*), and MATH-WRAT and MATH-TSMATH (*MATH)*. For ease of reading, the standardized measured variables are formatted in small caps, preceded by construct abbreviations, and the composites are formatted in large caps. Table 6 displays a full list of the variables included in the analyses, including their means, standard deviations, and ranges.

Missing data patterns were analyzed, resulting in either overall elimination from analysis (three participants who were missing data on both tasks within the same construct were eliminated from the analyses) or in missing value computations. Values were computed in seven instances: two instances of missing data from SS-FIGURES, two instances of missing WM-TTT data, one missing WM-DIGITS, one missing INH-SSRT, and one missing SS-SS-SW. Rather than imputing the mean for a given variable into the missing data fields, imputing a zero, or retaining the missing field, regression equations were generated to estimate missing

Table 6.

Descriptive Statistics for Unstandardized Variables (N = 148)

Variable	*M*	*SD*	*Range*
AGE	13.10	6.63	6.62 – 39.54
General Cognitive Ability			
GCA-BPVS	10.21	1.36	6.86 – 12.73
GCA-RAVENS	34.39	12.78	10 – 59
Working Memory Updating			
WM-DIGITS	23.41	4.817	14 – 35
WM-TTT	$1.50E^{3}$	$7.18E^{4}$	$2.07E^{5}$ – $3.37E^{3}$
Inhibition			
INH-SS-INH	31.99	4.36	18.55 – 41.88
INH-SSRT	2.88	.16	2.54 – 3.23
Set Shifting			
SS-SS-SW	.58	.28	-0.19 – 1.33
SS-FIGURES	288.01	31.92	185.46 – 353.03
Arithmetic			
MATH-WRAT	5.43	.90	3.00 – 7.28
MATH-TSMATH	3.37	.31	2.33 – 4.14

Note. GCA-BPVS and MATH-WRAT transformed using square root transformations. INH-SS-INH, INH-SSRT, SS-FIGURES, and MATH-TSMATH transformed using logarithmic transformations. INH-SSRT and MATH-TSMATH are displayed in their mirrored forms.

values per each variable from another variable within the same theoretical construct. This technique allowed for the missing value estimates to be closer to what might be expected for individual participants rather than using the mean for all participants, which might mask individual differences. In these instances, values were computed using simple regression and the resulting unstandardized regression estimates were inserted into the missing fields (see Appendix E). As a result of these calculations, there were no values missing from any of the variables included in the subsequent analyses.

While preparing the data for analysis, it was discovered that several variables violated one or more statistical assumption for parametric analyses. Namely, the variables of AGE, INH-SSRT, and MATH-TSMATH [1] displayed skewness and/or kurtosis values greater than 1.00. In addition, the variables of AGE, GCA-BPVS, GCA-RAVENS, INH-SS-INH, INH-SSRT, SS-FIGURES, MATH-TSMATH, and MATH-WRAT all appeared to be non-normally distributed according to histograms and normality tests. Thus, both logarithmic and square root transformations were performed on each of these variables (refer to recommended transformations in Field, 2009) with the exception of AGE, as the greater degree of variance within AGE was important to the theoretical basis of this study. Square root transformations were found to be effective in correcting the issue of non-normality for GCA-BPVS and MATH-WRAT. Logarithmic transformations were found to be effective in correcting non-normality on the RT-based variables: INH-SS-INH, INH-SSRT, SS-FIGURES, and MATH-TSMATH. In addition, the transformed INH-SSRT and MATH-TSMATH variables – both RT-based measures – were mirrored by multiplying them by -1, because increases in raw RT

1 Additionally, an overall accuracy variable was taken from the Tic-Tac-Toe measure and an efficiency variable from the Task Switch Math measure was computed and both were found to violate the assumption of normality; however, transformations were ineffective, thus these variables were not used in the analyses.

reflect *worse* rather than *better* performance. With the other variables within the two constructs associated with INH-SSRT and MATH-TSMATH (INH-SS-INH and MATH-WRAT, respectively), increasing values reflect better performance; thus, in order for the INH and MATH composites to be meaningful, the INH-SSRT and MATH-TSMATH variables were mirrored so that increasing values would reflect better performance in these variables, as well.

After cleaning and transforming the data, *z*-scores were formed for each of the variables to be included in the composite variable for each construct. Figure 7 displays age group means and standard errors for all standardized variables and Table 7 displays the results of separate analyses of variance (ANOVAs) examining age group difference within each variable. As all ANOVAs were significant, the results from Tukey's HSD post-hoc tests between age groups for each measure are provided in Table 8. The standardized, or centered, variables were formed on the basis of the means and standard deviations for each variable. The purpose of standardizing them was to allow for each variable to be on the same scale of measurement, so as to allow for them to be combined into a meaningful aggregate score. Statistical assumptions of linearity, normality, and homogeneity of variance were checked for WM, INH, SHIFT, and MATH and deemed adequate. GCA was found to be non-normally distributed; though further transformations were not possible for the composite (as one of the variables that went into forming the composite had already been transformed). The implications will be discussed in the path analysis description.

Zero-Order Correlations

Tables 9 and 10 display parametric and non-parametric correlations between AGE and all individual variables; Table 11 displays descriptive statistics and non-parametric

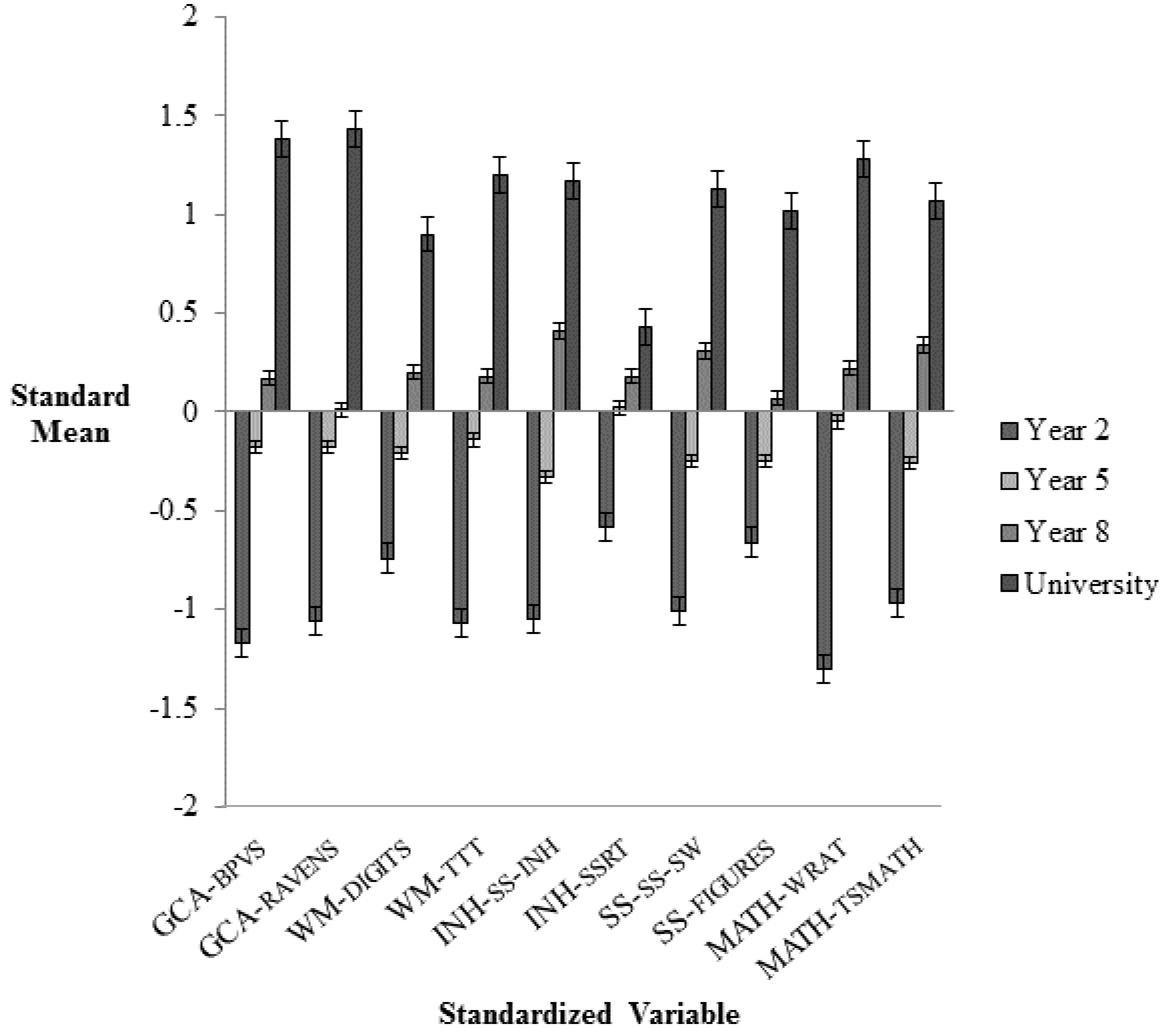


Figure 7. Average performance by age group on all standardized experimental variables. Vertical bars represent standard errors of the means.

Table 7

One-Way Analysis of Variance (ANOVA) for Age Group Effects across Standardized Variables (N = 148)

Variable	*SS*	*MS*	*F*(3, 144)	*p*	η^2
GCA-BPVS	112.17	37.39	154.60	< .001	.76
GCA-RAVENS	107.26	35.76	129.57	< .001	.73
WM-DIGITS	49.33	16.44	24.25	< .001	.34
WM-TTT	89.86	29.95	75.48	< .001	.61
INH-SS-INH	94.86	31.62	87.33	< .001	.65
INH-SSRT	19.30	6.44	7.36	< .001	.13
SS-SS-SW	84.17	28.06	64.31	< .001	.57
SS-FIGURES	52.09	17.36	26.34	< .001	.35
MATH-WRAT	115.44	38.48	175.56	< .001	.79
MATH-TSMATH	77.94	25.98	54.17	< .001	.53

Table 8.

Tukey's HSD Post-Hoc Test Results of Between-Group Comparisons Across Standardized Measures (N = 148)

Measure	Significant Pairs	*p*	Non-Significant Pairs	*p*
WM-DIGITS	Year-2 – Year-5	.025	Year-5 – Year-8	0.11
	Year-2 – Year-8	<.001		
	Year-2 – University	<.001		
	Year-5 – University	<.001		
	Year-8 – University	.004		
WM-TTT	Year-2 – Year-5	<.001	Year-5 – Year-8	.110
	Year-2 – Year-8	<.001		
	Year-2 – University	<.001		
	Year-5 – University	<.001		
	Year-8 – University	<.001		
INH-SS-INH	Year-2 – Year-5	<.001	None	--
	Year-2 – Year-8	<.001		
	Year-2 – University	<.001		
	Year-5 – Year-8	<.001		
	Year-5 – University	<.001		
	Year-8 – University	<.001		
INH-SSRT	Year-2 – Year-5	.026	Year-5 – Year-8	.871
	Year-2 – Year-8	.004	Year-5 – University	.243
	Year-2 – University	<.001	Year-8 – University	.698
SS-SS-SW	Year-2 – Year-5	<.001	None	--
	Year-2 – Year-8	<.001		
	Year-2 – University	<.001		
	Year-5 – Year-8	.001		
	Year-5 – University	<.001		
	Year-8 – University	<.001		
SS-FIGURES	Year-2 – Year-8	.001	Year-2 – Year-5	.118
	Year-2 – University	<.001	Year-5 – Year-8	.290
	Year-5 – University	<.001		
	Year-8 – University	<.001		
GCA-BPVS	Year-2 – Year-5	<.001	None	--
	Year-2 – Year-8	<.001		
	Year-2 – University	<.001		
	Year-5 – Year-8	.010		
	Year-5 – University	<.001		
	Year-8 – University	<.001		

Measure	Significant Pairs	*p*	Non-Significant Pairs	*p*
GCA-RAVENS	Year-2 – Year-5	<.001	Year-5 – Year-8	.353
	Year-2 – Year-8	<.001		
	Year-2 – University	<.001		
	Year-5 – University	<.001		
	Year-8 – University	<.001		
MATH-WRAT	Year-2 – Year-5	<.001	Year-5 – Year-8	.066
	Year-2 – Year-8	<.001		
	Year-2 – University	<.001		
	Year-5 – University	<.001		
	Year-8 – University	<.001		
MATH-TSMATH	Year-2 – Year-5	<.001	None	--
	Year-2 – Year-8	<.001		
	Year-2 – University	<.001		
	Year-5 – Year-8	.001		
	Year-5 – University	<.001		
	Year-8 – University	<.001		

Table 9.

Parametric Zero-Order Correlations for Unstandardized Variables, Grouped by Composite (N = 148)

Variable	1	2	3	4	5	6	7	8	9	10	11
1. AGE	--										
General Cognitive Ability											
2. GCA-BPVS	.805	--									
3. GCA-RAVENS	.781	.835	--								
Working Memory Updating											
4. WM-DIGITS	.583	.628	.578	--							
5. WM-TTT	.690	.761	.705	.484	--						
Inhibition											
6. INH-SS-INH	.690	.721	.662	.563	.685	--					
7. INH-SSRT	.218	.345	.301	.176	.336	.324	--				
Set Shifting											
8. SS-SS-SW	.656	.711	.732	.534	.682	.730	.272	--			
9. SS-FIGURES	.594	.573	.574	.334	.546	.548	.288	.586	--		
Arithmetic											
10. MATH-WRAT	.753	.856	.831	.576	.777	.752	.372	.785	.567	--	
11. MATH-TSMATH	.645	.594	.618	.366	.585	.636	.248	.609	.425	.697	--

Note. Correlations greater than .17 are significant at $p < .05$; correlations greater than .29 are significant at $p < .01$. Parametric correlations based on Pearson's r.

Table 10.

Non-Parametric Zero-Order Correlations for Unstandardized Variables, Grouped by Composite (N = 148)

Variable	1	2	3	4	5	6	7	8	9	10	11
1. AGE	--										
General Cognitive Ability											
2. GCA-BPVS	.846	--									
3. GCA-RAVENS	.777	.787	--								
Working Memory Updating											
4. WM-DIGITS	.558	.624	.535	--							
5. WM-TTT	.767	.739	.686	.450	--						
Inhibition											
6. INH-SS-INH	.793	.721	.650	.532	.692	--					
7. INH-SSRT	.357	.340	.296	.155	.345	.324	--				
Set Shifting											
8. SS-SS-SW	.741	.711	.716	.518	.685	.736	.271	--			
9. SS-FIGURES	.620	.615	.598	.344	.602	.592	.312	.579	--		
Arithmetic											
10. MATH-WRAT	.822	.841	.821	.549	.785	.764	.370	.819	.632	--	
11. MATH-TSMATH	.791	.694	.678	.399	.676	.719	.283	.678	.524	.784	--

Note. Correlations greater than .17 are significant at $p < .05$; correlations greater than .29 are significant at $p < .01$. Non-parametric correlations based on Spearman's r_s.

Table 11.

Descriptive Statistics and Non-Parametric Zero-Order Correlations for Composite and Standardized Variables Included in Path Models (N = 148)

	Variable	*Min*	*Max*	*SD*	1	2	3	4	5	6	7	8
1.	AGE	6.62	39.54	6.63	--	.828	.739	.558	.702	.753	.645	.759
2.	GCA	-2.15	1.85	.96		--	.810	.651	.759	.881	.633	.821
3.	WM	-1.80	2.19	.86			--	.627	.683	.786	.552	.726
4.	INH	-2.36	1.80	.81				--	.634	.691	.544	.670
5.	SHIFT	-2.70	2.15	.89					--	.764	.581	.727
6.	MATH-WRAT	-2.71	2.06	1.00						--	.697	.921
7.	MATH-TSMATH	-2.52	3.40	1.00							--	.921
8.	MATH	-2.12	1.87	.92								--

Note. Correlations greater than .17 are significant at $p < .05$; correlations greater than .29 are significant at $p < .01$. Parametric correlations based on Pearson's r, non-parametric correlations based on Spearman's r_s. Means for all composite and standardized variables = 0.

correlations between AGE and composite variables. As AGE and GCA were found to violate the statistical assumption of normality, the following description will focus on Spearman's non-parametric correlation coefficients rather than Pearson's correlation coefficients as Spearman's provides more conservative correlation estimates. To illustrate how these coefficients differed depending on method, both coefficient values are provided in Tables A and B.

Correlations within composites. All correlations within composites were significant at $p < .01$, and correlations were strongest between GCA ($r_s = .79$, $p < .001$), MATH ($r_s = .78$, $p < .001$), and SHIFT ($r_s = .58$, $p < .001$). The high correlation between MATH-WRAT and MATH-TSMATH indicates that proficiency and fluency in arithmetic are related constructs. Although the correlation between the INH-SS-INH and INH-SSRT was slightly smaller ($r_s = .32$, $p < .001$), the integrity of INH is not likely to be affected given that INH-SSRT was not highly correlated with any one other variable (the highest correlation was between INH-SSRT and MATH-WRAT, $r_s = .37$, $p < .001$). Similar to the variables in INH, the correlation between the variables comprising WM was moderate ($r_s = .45$, $p < .001$). In considering the integrity of the composite variables, it is important to keep in mind that the correlations between both variables within INH and WM were significant at $p < .001$.

Correlations between age and experimental variables. On the level of both individual variables and composite variables, AGE was positively correlated with all experimental variables ($p < .001$). The highest correlations were found between AGE and GCA ($r_s = .85$, $p < .001$) and between AGE and MATH ($r_s = .84$, $p < .001$). Thus, while older participants demonstrated greater performance across all constructs, these differences were largest in terms of performance on measures of GCA and MATH. On an individual variable basis, the lowest correlation (though still significant) was found between AGE and INH-SSRT (r_s = .36, p < .001). The high correlation between AGE and INH (composite; r_s = .71, p < .001) may be reflective of the high correlation between AGE and INH-SS-INH performance (r_s = .79, p < .001). Thus, in this sample, age affected the aspects of inhibition that were measured by the Shape School task to a larger degree than the aspects of inhibition that were measured by the Stop-Signal task.

Correlations between arithmetic and experimental variables. All correlations between experimental variables and MATH-WRAT were significant ($p \leq .001$) and all correlations between experimental variables and MATH-TSMATH were significant ($p \leq .001$), indicating that the experimental variables significantly correlated with both proficiency and fluency in arithmetic. In addition, all correlations between experimental composites and MATH were significant (p < .001). As all correlations were positive, the data indicate that there is a positive relation between GCA, WM, INH, and SHIFT and arithmetic performance such that as performance in one of these cognitive domains improves, arithmetic performance improves as well.

Correlations separated by age group. Separate non-parametric correlation analyses that included age, standardized predictor variables, and standardized arithmetic variables were conducted for each of the four age groups in this study (see Tables 12, 13, 14, and 15). The pattern of correlations between experimental variables and arithmetic variables was inconsistent across age groups in several instances.

First, though AGE was not significantly related to either of the arithmetic variables among Year-2, Year-5, and adult participants, it was significantly negatively related to MATH-WRAT performance among Year-8 participants. Along with the negative – though

Table 12.

Non-parametric Correlations for Standardized Variables Across Year-2 Participants (n = 36)

Variable	1	2	3	4	5	6	7	8	9	10	11
12. AGE	--										
General Cognitive Ability											
13. GCA-BPVS	.332	--									
14. GCA-RAVENS	.253	.342	--								
Working Memory Updating											
15. WM-DIGITS	.235	.252	.308	--							
16. WM-TTT	.226	.388	.240	.253	--						
Inhibition											
17. INH-SS-INH	.250	.340	.225	.360	.134	--					
18. INH-SSRT	.287	.298	.344	.076	.294	.170	--				
Set Shifting											
19. SS-SS-SW	-.005	.282	.256	.156	.223	.254	.155	--			
20. SS-FIGURES	-.072	.129	.321	.234	.171	-.083	.079	.200	--		
Arithmetic											
21. MATH-WRAT	.167	.495	.320	.409	.225	.306	.450	.382	.088	--	
22. MATH-TSMATH	.112	-.130	.021	-.122	-.168	.077	-.120	.018	-.163	.097	--

Note. Correlations > .330 are significant at $p < .05$; > .430 significant at $p < .01$. Correlations based on Spearman's r_s. AGE not standardized. Standardization based on total sample.

Table 13.

Non-parametric Correlations for Standardized Variables Across Year-5 Participants (n = 44)

Variable	1	2	3	4	5	6	7	8	9	10	11
1. AGE	--										
General Cognitive Ability											
2. GCA-BPVS	.162	--									
3. GCA-RAVENS	.145	.049	--								
Working Memory Updating											
4. WM-DIGITS	-.117	.337	.119	--							
5. WM-TTT	.412	.099	.064	-.243	--						
Inhibition											
6. INH-SS-INH	.049	.131	.007	.022	.339	--					
7. INH-SSRT	.274	.181	-.046	-.039	.172	.186	--				
Set Shifting											
8. SS-SS-SW	.078	.148	.122	.195	.302	.390	-.063	--			
9. SS-FIGURES	.184	.242	.082	-.133	.358	.162	.329	-.035	--		
Arithmetic											
10. MATH-WRAT	.043	.391	.453	.232	.342	.435	.070	.483	.209	--	
11. MATH-TSMATH	.263	.117	.253	-.027	.598	.417	.218	.361	.170	.657	--

Note. Correlations > .290 are significant at $p < .05$; > .360 significant at $p < .01$; > .580 significant at $p < .001$. Correlations based on Spearman's r_s. AGE not standardized. Standardization based on total sample.

Table 14.

Non-parametric Correlations for Standardized Variables Across Year-8 Participants (n = 36)

Variable	1	2	3	4	5	6	7	8	9	10	11
1. AGE	--										
General Cognitive Ability											
2. GCA-BPVS	.037	--									
3. GCA-RAVENS	-.205	.402	--								
Working Memory Updating											
4. WM-DIGITS	.270	.302	-.021	--							
5. WM-TTT	-.181	.387	.166	.212	--						
Inhibition											
6. INH-SS-INH	-.046	.088	-.013	.103	.179	--					
7. INH-SSRT	-.084	.048	-.061	-.268	-.008	.181	--				
Set Shifting											
8. SS-SS-SW	-.241	.020	.499	-.026	.027	.287	-.005	--			
9. SS-FIGURES	-.148	.237	.199	-.093	.064	.325	.294	.371	--		
Arithmetic											
10. MATH-WRAT	-.342	.274	.358	.072	.228	.203	.042	.471	.412	--	
11. MATH-TSMATH	-.023	.353	.198	.070	.172	.240	.256	.316	.218	.499	--

Note. Correlations > .330 are significant at $p < .05$; > .460 significant at $p < .01$. AGE not standardized. Correlations based on Spearman's r_s. Standardization based on total sample.

Table 15.

Non-parametric Correlations for Standardized Variables Across Adult Participants (n = 32)

Variable	1	2	3	4	5	6	7	8	9	10	11
1. AGE	--										
General Cognitive Ability											
2. GCA-BPVS	.319	--									
3. GCA-RAVENS	.003	.387	--								
Working Memory Updating											
4. WM-DIGITS	.333	.421	.281	--							
5. WM-TTT	.029	.329	.112	.178	--						
Inhibition											
6. INH-SS-INH	-.200	.033	.107	.321	.132	--					
7. INH-SSRT	-.391	-.309	-.059	-.010	.120	-.120	--				
Set Shifting											
8. SS-SS-SW	-.046	.307	.357	.331	.369	.416	.013	--			
9. SS-FIGURES	.269	.245	-.035	.341	-.137	.445	-.212	.350	--		
Arithmetic											
10. MATH-WRAT	.009	.386	.400	.133	.207	.109	-.100	.443	.209	--	
11. MATH-TSMATH	.150	.441	.438	.142	.164	.146	-.200	.411	.173	.762	--

Note. Correlations > .350 are significant at $p < .05$; > .750 significant at $p < .001$. Correlations based on Spearman's r_s. AGE not standardized. Standardization based on total sample.

non-significant – relation between AGE and MATH-TSMATH scores within this age group, Year-8 was the only group to which older participants demonstrated worse performance on measures of arithmetic. Also, the relation between GCA-BPVS and MATH-TSMATH scores only was significant in the positive correlations found in the Year-8 and adult age groups. The negative non-significant correlation between these two variables in the Year-2 age group indicates that the direction of the relation between GCA-BPVS and MATH-TSMATH performance differed depending on age group. Though correlations between GCA-RAVENS and MATH-TSMATH were positive across all age groups, the only significant correlation was in the adult age group. Thus, the direction and magnitude of the relations between arithmetic and age and arithmetic and general cognitive ability varied across age groups.

Similar to the pattern found between GCA-RAVENS and MATH-TSMATH, although all correlations between WM-DIGITS and MATH-WRAT were positive, only one age group correlation was significant, i.e., Year-2. None of the correlations between WM-DIGITS and MATH-TSMATH across the four age groups were significant. Positive correlations between WM-TTT and MATH-WRAT and between WM-TTT and MATH-TSMATH were only significant for Year-5 participants. All non-significant correlations between WM-TTT and MATH-WRAT also were positive, and the non-significant correlations between WM-TTT and MATH-TSMATH in the Year-8 and adult age groups were positive while this correlation was negative among Year-2 participants. Therefore, the magnitude of the relation between WM-DIGITS and arithmetic and between WM-TTT and arithmetic varied by age group and type of arithmetic measured, and the direction of the relation between WM-TTT and arithmetic varied – though not significantly – depending on age group.

Year-5 was the only group for which INH-SS-INH scores were significantly correlated with either arithmetic variable. Though none of the other groups' INH-SS-INH scores were significantly correlated with MATH-WRAT or MATH-TSMATH, correlations between these two arithmetic variables and INH-SS-INH were positive and significant for Year-5. Additionally, correlations between the second inhibition variable, INH-SSRT, and arithmetic variables were not consistent across age groups. The only significant correlation between INH-SSRT and either arithmetic variable was between INH-SSRT and MATH-WRAT scores among Year-2 participants. While the correlations between INH-SSRT and MATH-WRAT were positive for Year-2, Year-5, and Year-8 participants, INH-SSRT was negatively (non-significantly) correlated with MATH-WRAT among adults. None of the correlations between INH-SSRT and MATH-TSMATH were significant, and while they were all similar in magnitude, half were negative and half were positive. The non-significant correlations between INH-SSRT and MATH-TSMATH were positive among the middle two age groups, i.e., Year-5 and Year-8, and negative among the youngest, i.e., Year-2, and oldest, i.e., adults, age groups. These results indicate that the magnitude of the relation between the aspect(s) of inhibition measured by the Shape School task and arithmetic depended on age group; while both the magnitude and the direction of the relation between the aspect(s) of inhibition measured by the Stop-Signal task and arithmetic depended on age group.

Correlations between SS-SS-SW and MATH-WRAT scores were positive and significant across all age groups. Thus, better performance on the Switch Condition of the Shape School task predicted better performance on the MATH-WRAT across all age groups. Correlations between SS-SS-SW and MATH-TSMATH scores also were positive across age groups, and were significant for Year-5 and adult participants (but not for Year-2 or Year-8 participants). Similarly, all correlations between MATH-WRAT and the second measure of set shifting, SS-FIGURES, were positive, though only one (Year-8) was significant. On the contrary, none of the correlations between MATH-TSMATH and SS-FIGURES were significant and not all were positive. Although the non-significant correlations between MATH-TSMATH and SS-FIGURES were positive for Year-5, Year-8, and adult participants, this correlation was negative for Year-2 participants. Therefore, arithmetic was at least partially related to the aspect of set shifting that was tapped by the Shape School task, though, in comparison, it was largely unrelated to the aspect of set shifting that was tapped by the Figure Matching task.

Most of the correlations between the two arithmetic variables were significant, and all were positive. The only non-significant correlation between MATH-WRAT and MATH-TSMATH was in the youngest (Year-2) age group. These results indicate that the relation between the two arithmetic measures was mostly consistent across age groups. In sum, age, general cognitive ability, and executive functions differentially related arithmetic performance based on age group, measure used, and form of arithmetic measured.

Regression Analyses

Seven sets of regression analyses were conducted to determine the unique and combined influences of age, general cognitive ability, working memory updating, inhibition, and set shifting on overall arithmetic performance, arithmetic proficiency, and arithmetic fluency. Each set contained three separate analyses that included the same independent variables (IVs) and the same steps, but each involved different dependent variables (DVs). The three DVs reflected different measures of arithmetic: arithmetic proficiency (MATH-WRAT), arithmetic fluency (MATH-TSMATH), and overall performance (MATH). As regression assumes normality of the residuals (Field, 2009), normal probability plots and histograms of the standardized residuals produced from the regression analyses were checked and the assumptions of normality and homogeneity of variance did not appear to be violated.

Set 1: Two-Step models. These models are depicted in Tables 16, 17, and 18. In these models, Step 1 included AGE and the variables associated with general cognitive ability, namely GCA-BPVS and GCA-RAVENS, and Step 2 included all variables associated with executive function, i.e., WM-DIGITS, WM-TTT, INH-SS-INH, INH-SSRT, SS-SS-SW, and SS-FIGURES. In the MATH composite model, AGE and GCA-RAVENS continued to account for a significant portion of variance in math performance in Step 2 once the influence of the executive function variables had been added to the model, though GCA-BPVS no longer remained significant (see Table 16). In this model, SS-INH accounted for 20% of the variance in MATH performance and SS-SS-SW accounted for 19% and were the only executive function variables to account for a significant portion of variance in MATH scores once all other variables had been entered (at $p < .05$). All together, the variables included in Step 1 of this model accounted for 69.5% of the variance in MATH scores, F (3, 144) = 109.20, $p < .001$, $\Delta R^2 = .70$, with the executive function variables added in Step 2 contributing an additional 7.3% of variance in the model (a total of 76.8% of variance accounted for in the final model), F (9, 138) = 50.79, $p < .001$, $\Delta R^2 = .07$. Thus, inhibition and set shifting continued to account for overall arithmetic performance even after the significant effects of age and general cognitive ability taken into account. However, other variables related to inhibition and set shifting, as well as variables related to working memory updating, did not account for overall arithmetic performance.

The results from the MATH-TSMATH model provide an explanation for the discrepancies found between the MATH composite model and the MATH-WRAT model. In this model, AGE accounted for a significant portion of arithmetic fluency scores when entered

Table 16.

MATH Composite Model 1: Regression Analysis Summary for Standardized Variables Predicting Composite Arithmetic Performance (N = 148)

	Step 1					Step 2				
Variable	*B*	*SEB*	*ß*	*t*	*p*	*B*	*SEB*	*ß*	*t*	*p*
AGE	.04	.01	.25	3.06	.003	.03	.01	.20	2.48	.014
GCA-BPVS	.28	.09	.31	3.26	.001	.11	.09	.12	1.26	.209
GCA-RAVENS	.31	.08	.34	3.78	.000	.21	.08	.22	2.65	.009
WM-DIGITS						-.08	.05	-.09	-1.56	.120
WM-TTT						.12	.06	.13	1.90	.059
INH-SS-INH						.20	.07	.21	3.05	.003
INH-SSRT						.05	.04	.05	1.14	.254
SS-SS-SW						.19	.07	.21	2.94	.004
SS-FIGURES						-.07	.05	-.07	-1.29	.200
R^2	.70					.77				
Adj R^2	.69					.75				
F(df)	109.20 (3, 144)				.000					
ΔF(df)						7.29 (6. 138)				.000

Table 17.

MATH-WRAT Model 1: Regression Analysis Summary for Standardized Variables Predicting MATH-WRAT Performance (N = 148)

	Step 1					Step 2				
Variable	*B*	*SEB*	*ß*	*t*	*p*	*B*	*SEB*	*ß*	*t*	*p*
AGE	.01	.01	.07	1.00	.319	.00	.01	.01	.22	.829
GCA-BPVS	.50	.08	.50	6.32	.000	.31	.08	.31	3.90	.000
GCA-RAVENS	.36	.08	.36	4.73	.000	.24	.07	.24	3.31	.001
WM-DIGITS						-.01	.05	-.01	-.18	.861
WM-TTT						.15	.06	.15	2.55	.012
INH-SS-INH						.12	.06	.12	1.93	.056
INH-SSRT						.06	.04	.06	1.55	.124
SS-SS-SW						.21	.06	.21	3.39	.001
SS-FIGURES						-.03	.05	-.03	-.72	.475
R^2	.78					.83				
Adj R^2	.77					.82				
F(df)	169.26 (3, 144)				.000					
Δ*F*(df)						7.51 (6, 138)				.000

Table 18.

MATH-TSMATH Model 1: Regression Analysis Summary for Standardized Variables Predicting MATH-TSMATH Performance (N = 148)

	Step 1					Step 2				
Variable	*B*	*SEB*	*ß*	*t*	*p*	*B*	*SEB*	*ß*	*t*	*p*
AGE	.06	.02	.39	3.57	.000	.05	.02	.35	3.09	.002
GCA-BPVS	.06	.13	.06	.48	.633	-.09	.13	-.09	-.69	.492
GCA-RAVENS	.26	.12	.26	2.19	.030	.18	.12	.18	1.46	.146
WM-DIGITS						-.15	.08	-.15	-1.92	.057
WM-TTT						.09	.10	.09	.94	.348
INH-SS-INH						.28	.10	.28	2.80	.006
INH-SSRT						.04	.06	.04	.56	.576
SS-SS-SW						.18	.10	.18	1.78	.077
SS-FIGURES						-.10	.08	-.10	-1.24	.217
R^2	.45					.53				
Adj R^2	.44					.50				
F(df)	39.40 (3, 144)				.000					
ΔF(df)						3.91 (6, 138)				.001

with GCA-BPVS and GCA-RAVENS in Step 1, and continued to significantly account for performance when entered with the remaining variables in Step 2 (see Table 18). Thus, MATH-TSMATH performance, i.e., arithmetic fluency, seems to be driving the relation between AGE and MATH scores. In the MATH-TSMATH model, the only executive function variable that significantly accounted for arithmetic fluency scores was INH-SS-INH, accounting for 27.9% of the variance in MATH-TSMATH scores. As INH-SS-INH did not significantly account for MATH-WRAT scores, the significant relation between overall arithmetic performance (MATH scores) and INH-SS-INH also may have been due to the influence of MATH-TSMATH in the MATH composite. However, unlike the MATH model, the significant portion of variance that was accounted for by GCA-RAVENS scores in Step 1 of the MATH-TSMATH model was no longer significant when entered with executive function variables in Step 2. Therefore, the significant relation between MATH scores and GCA-RAVENS scores can be accounted for by the significant relation between MATH-WRAT scores, i.e., arithmetic proficiency, and GCA-RAVENS scores. In total, AGE, GCA-BPVS, and GCA-RAVENS accounted for 45.1% of the variance in MATH-TSMATH scores, F (3, 144) = 39.40, $p < .001$, $\Delta R^2 = .45$, while the executive function variables added in Step 2 accounted for an additional 8.0%, F (9, 138) = 17.34, $p < .001$, $\Delta R^2 = .08$.

Sets 2-7: Five-Step models. To identify the exact contribution of each construct to the arithmetic performance, six sets of three hierarchical regressions were conducted. Within each set, the variables included and the steps in which they were included were the same, but the DV differed for each model, i.e., MATH, MATH-WRAT, or MATH-TSMATH. Across all models, AGE and GCA-BPVS/GCA-RAVENS were entered as Steps 1 and 2, respectively. In the first set (Set 2), the variables associated with working memory updating were entered in Step 3, the inhibition variables were entered in Step 4, and the set shifting variables were entered in Step 5 (see Table 19). In the next set (Set 3), the ordering of inhibition and set shifting was reversed. As the pattern of results was consistent across sets, only the first two sets will be discussed; however, the reader may refer to Appendix F for a summary of the remaining 14 models.

Table 19.

*Summary of Select Five-Step Hierarchical Regression Analyses for Standardized Variables Predicting MATH Composite, MATH-*WRAT*, and MATH-*TSMATH *Scores (N = 148)*

Variables Added	R^2	Adj R^2	ΔR^2	ΔF(df)	*p*
MATH Composite Model 2					
Step 1: AGE	.58	.57	.58	198.00 (1, 146)	< .001
Step 2: GCA-RAVENS, GCA-BPVS	.70	.69	.12	28.08 (2, 144)	< .001
Step 3: WM-DIGITS, WM-TTT	.72	.71	.02	6.13 (2, 142)	.003
Step 4: INH-SS-INH, INH-SSRT	.75	.74	.03	9.57 (2, 140)	< .001
Step 5: SS-SS-SW, SS-FIGURES	.77	.75	.02	4.58 (2, 138)	.012
MATH-WRAT Model 2					
Step 1: AGE	.57	.56	.56	190.67 (1, 146)	< .001
Step 2: GCA-RAVENS, GCA-BPVS	.78	.77	.21	69.33 (2, 144)	< .001
Step 3: WM-DIGITS, WM-TTT	.80	.80	.02	8.43 (2, 142)	< .001
Step 4: INH-SS-INH, INH-SSRT	.82	.81	.02	6.62 (2, 140)	.002
Step 5: SS-SS-SW, SS-FIGURES	.83	.82	.01	5.76 (2, 138)	.004
MATH-TSMATH Model 2					
Step 1: AGE	.42	.41	.42	104.19 (1, 146)	< .001
Step 2: GCA-RAVENS, GCA-BPVS	.45	.44	.03	4.51 (2, 144)	.013
Step 3: WM-DIGITS, WM-TTT	.47	.45	.02	2.92 (2, 142)	.057
Step 4: INH-SS-INH, INH-SSRT	.52	.49	.05	6.46 (2, 140)	.002
Step 5: SS-SS-SW, SS-FIGURES	.53	.50	.01	2.01 (2, 138)	.139
MATH-TSMATH Model 3					
Step 1: AGE	.42	.41	.42	104.19 (1, 146)	< .001
Step 2: GCA-RAVENS, GCA-BPVS	.45	.44	.03	4.51 (2, 144)	.013
Step 3: WM-DIGITS, WM-TTT	.47	.45	.02	2.92 (2, 142)	.057
Step 4: SS-SS-SW, SS-FIGURES	.50	.48	.03	4.04 (2, 140)	.020
Step 5: INH-SS-INH, INH-SSRT	.53	.50	.03	4.33 (2, 138)	.015

In both the MATH model and the MATH-WRAT model of the first set, each step accounted for a significant portion of variance in arithmetic scores, indicating that AGE, GCA, working memory updating, inhibition, and set shifting each have a significant positive relation with overall arithmetic performance and arithmetic proficiency, specifically. In the MATH-WRAT model, age and GCA accounted for 77.9% of the variance in scores, working memory explained an additional 2.3%, inhibition accounted for 1.4% of the remaining variance, and set shifting accounted for a final 1.4% of variance; together, these variables accounted for 83.3% of the variance in MATH-WRAT scores. Unlike the MATH composite and MATH-WRAT models, the variables associated with working memory updating and set shifting did not account for a significant portion of variance in MATH-TSMATH skills when entered into the model in Steps 3 and 5, respectively. When entered in Steps 1 and 2, AGE and GCA-BPVS/GCA-RAVENS, respectively, accounted for 45.1% of the variance in MATH-TSMATH scores; working memory updating accounted for 2.2% of the remaining variance (non-significant). Entered in Step 4, inhibition accounted for a significant 4.5% of the remaining variance and set shifting accounted for 1.4% of the variance remaining in Step 5.

In Set 3, the results produced in the MATH composite model and the MATH-WRAT model were consistent with the results from the previous set; however, the results from the MATH-TSMATH model in Set 3 were slightly different from the results from Set 2. While the variables related to working memory remained non-significant when added in Step 3, the variables related to set shifting accounted for significant portions of variance in MATH-TSMATH scores when entered in Step 4 rather than Step 5. Across all sets of models, when the order of inhibition and set shifting was reversed, this same pattern of results emerged; set shifting was significant only when it preceded inhibition (see Appendix F). From a theoretical standpoint, these results are consistent with the common assumption that inhibition is required in order to successfully shift mental sets (e.g., Mayr, 2002; Meiran, 1996).

The part and partial correlations, provided in Table 20, were the same in the final step of the two-step models as they were in the final step of the five-step models. While the squared partial correlation value for each variable provides an estimate of the percentage of variance left-over in the outcome variable (here, MATH, MATH-WRAT, or MATH-TSMATH scores) that is accounted for by the target predictor variable after the variance from all other predictors has been accounted for, the squared semi-partial correlation value provides an estimate of the amount that R^2 would decrease if the target variable were to be removed from the prediction model (Cohen, Cohen, West, & Aiken, 2003). Squared partial correlations indicated that after the variance from all other variables had been accounted for, AGE accounted for less than 1% of the variance in MATH-WRAT scores, but 6.3% of variance in MATH-TSMATH scores and 4.4% of variance in overall MATH scores. In addition, taking AGE out of the model would reduce the R^2 in the MATH-TSMATH model from .53 to .50, though removing AGE from the MATH-WRAT model would not cause the R^2 value to change from .83. Thus, AGE accounted for more unique and shared variance in the MATH-TSMATH model than in the MATH-WRAT model.

Also of note, although GCA-BPVS contributed an additional 10.2% of variance in MATH-WRAT scores after all other variables had been accounted for, it contributed just .04% of additional variance unaccounted for in MATH-TSMATH scores. Additionally, though removing GCA-BPVS from the MATH-WRAT prediction model would reduce the R^2 value from .83 to .81, removing GCA-BPVS from the MATH-TSMATH model would not change the

Table 20.

Partial and Semi-Partial (Semi-Part) Correlations between Standardized IVs and DVs in the Final Step of Each Regression Model (N = 148)

	Composite Model		MATH-WRAT Model		MATH-TSMATH Model	
Variable	Partial	Semi-Part	Partial	Semi-Part	Partial	Semi-Part
AGE	.21	.10	.02	.01	.25	.18
GCA-BPVS	.11	.05	.32	.14	-.06	-.04
GCA-RAVENS	.22	.11	.27	.12	.12	.09
WM-DIGITS	-.13	-.06	-.02	-.01	-.16	-.11
WM-TTT	.16	.08	.21	.09	.08	.06
INH-SS-INH	.25	.13	.16	.07	.23	.16
INH-SSRT	.10	.05	.13	.05	.05	.03
SS-SS-SW	.24	.12	.28	.12	.15	.10
SS-FIGURES	-.11	-.05	-.06	-.03	-.11	-.07

R^2 value from .52. This pattern is similar to the pattern of partial and semi-partial correlations for WM-TTT and SS-SS-SW. While WM-TTT accounted for an additional 4.4% of the variance left-over in the MATH-WRAT model, it accounted for just .06% of the variance left-over in the MATH-TSMATH model. SS-SS-SW accounted for 7.8% of the left-over variance in MATH-WRAT scores but only 2.3% in the MATH-TSMATH model. INH-SS-INH displayed an opposite pattern; it accounted for 5.3% of the left-over variance in MATH-TSMATH scores and 2.6% of the left-over variance in MATH-WRAT scores. In addition, removing INH-SS-INH from the MATH-TSMATH model would reduce the R^2 value to .50 (from .53) but removing INH-SS-INH from the MATH-WRAT model would not change the R^2 value from .83.

Path Analyses

Path analyses were conducted using the non-parametric correlations in the correlation matrix provided in Table 11. Spearman's r_s correlation coefficients were used in place of Pearson's r based on the violated assumption of normality in two of the variables (AGE and GCA). The specific matrices are reported appropriately, i.e., beta (β) represents the path estimates between two endogenous variables, psi (ψ) represents correlations between endogenous variables, and gamma (γ) represents the path estimates between exogenous and endogenous variables.

Model Testing. Two sets of models were tested; in the first set, the endogenous arithmetic variable used was MATH, reflecting overall arithmetic performance. The models in this set will be referred to as the composite models. The full composite model is shown in Figure 6. The second set of models included the standardized MATH-TSMATH and MATH-WRAT endogenous variables as reflections of arithmetic proficiency and fluency, respectively. These models will be referred to as the combined models. The full combined model is identical to Figure 6 with the exception of an additional arithmetic endogenous variable. Models were fitted using the LISREL 8.80 Student Edition (Jöreskog & Sörbom, 2009), which generates parameter estimates using the maximum likelihood technique. Chi-square difference tests were used to compare several restricted models against the full, or saturated, models in order to find the most parsimonious models that minimized the differences between the predicted correlation matrices and the actual correlation matrices, i.e., models with minimal residual values (Jöreskog & Sörbom, 1989). Paths with small standardized coefficient values were systematically removed from the full models and the resulting nested models were compared against the previous more saturated models to determine whether the excluded paths significantly improved or worsened the predictability of the model.

Composite model. In the full composite model, there was one non-significant path from WM to MATH ($\beta = .05$) and one non-significant correlation between WM and INH ($\psi = .06$). In addition, the fit indices suggested that the model was not a good fit for the data, given that the *p*-value associated with the χ^2 of the model was significant at $p < .001$ and the root mean square error of approximation (RMSEA) of the model exceeded the acceptable range of $\leq .05$ to .08 (Kline, 2005 recommends using χ^2, RMSEA, the 90% confidence interval for RMSEA, CFI, and SRMR to judge model fit). Additional goodness of fit indices are provided in Table 21. The non-significant paths were removed from the model, resulting in a nested model (which will be referred to as the "modified model") that was tested against the full model using a χ^2 difference test, $\Delta\chi^2 = 3.40$, $\Delta df = 2$, $p = .18$. After determining that these paths did not significantly improve the fit of the model, the full model was rejected in favor of the modified model.

The modified composite model gained two additional degrees of freedom, making it possible to introduce three new correlations to the model linking GCA to each of the executive function variables. These paths were not included in the original model because including them would have caused the model to be "just-identified" or "saturated," meaning the number of known values would equal the number of unknown values in the model and only one unique solution would be possible, rather than multiple solutions that could be tested against one another as is the case in over-identified models (Kelloway, 1998). Of the three new correlations added to the model, one was non-significant – the correlation between GCA and INH ($\psi = .01$) – and was removed from the model. A χ^2 difference test revealed that this correlation did not significantly contribute to the predictability of the model, $\Delta\chi^2 = 0.22$, $\Delta df = 1$, $p = .64$; thus the nested model was retained (referred to as the "final model"). All paths included in the final composite model were found to be significant at $p < .05$, and goodness of fit indices indicated that the model provided a good fit to the data (refer to Table 21).

Combined model. A full combined model was tested in order to identify the path structure associated with both proficiency and fluency in arithmetic, as the regression analyses indicated that the predictor variables contributed to each of these outcome variables differently. To remain consistent across model sets, the full combined model included paths from AGE to each of the six experimental variables, from GCA, WM, INH, and SHIFT to the two arithmetic variables (MATH-TSMATH and MATH-WRAT), as well as three correlations linking each of the executive function variables to each other. The goodness of fit statistics (displayed in Table 21) indicated that this model may not have fit the data well, with a significant *p*-value associated with the χ^2 and high RMSEA value. In addition, examination of the path coefficients indicated that several paths may not have been needed in the model.

Results from the full model revealed six non-significant paths: (1) from GCA to MATH-TSMATH (β = .14), (2) from WM to MATH-TSMATH (β = -.04), (3) from INH to MATH-TSMATH (β = .08), (4) from SHIFT to MATH-TSMATH (β = .13), (5) from AGE to MATH-WRAT (γ = .09), and the correlation linking WM and INH (ψ = .06). All non-significant paths were removed one at a time and, based on a χ^2 difference test, it was discovered that the model that removed all non-significant paths except for the path from SHIFT to MATH-TSMATH, i.e., the slightly more saturated model, was a better fit for the data than the model that removed all six non-significant paths, i.e., the more parsimonious model, $\Delta\chi^2 = 5.47$, $\Delta df = 1$, $p = .02$. Thus, the decision was to retain the path from SHIFT to MATH-TSMATH but to remove all other non-significant paths. This model is referred to as the "modified model"; the goodness of fit statistics are provided in Table 21.

Table 21.

Goodness of Fit Indices for Theoretical, Modified, and Final Path Models (N = 148)

Model	χ^2	df	*p*	*RMSEA*	*RMSEA 90% C.I.*	*CFI*	*SRMR*
MATH Composite							
Full	39.79	3	<.001	0.27	0.22 – 0.32	0.97	0.06
Modified	43.19	5	<.001	0.22	0.16 – 0.28	0.97	0.05
Final	3.61	3	.31	0.04	0.00 – 0.15	1.00	0.02
MATH-TSMATH and MATH-WRAT Combined							
Full	57.89	4	<.001	0.29	0.22 – 0.36	0.97	0.05
Modified	64.86	9	<.001	0.19	0.15 – 0.24	0.97	0.05
Final	25.29	7	<.001	0.13	0.07 – 0.18	0.99	0.03

As with the composite model, three correlations were added to the modified combined model linking GCA with each of the variables associated with executive function. This model was compared against the modified model and it was determined that the more parsimonious modified model fit the data significantly worse than the more saturated model that contained these correlations, $\Delta\chi^2 = 39.79$, $\Delta df = 3$, $p < .001$. However, the more saturated model contained one non-significant path – the correlation between GCA and INH ($\psi = .01$) – and removing this path did not significantly affect the fit of the model, $\Delta\chi^2 = 0.22$, $\Delta df = 1$, $p = .64$. Thus, the more parsimonious model was retained and will be referred to as the "final model." The goodness-of-fit indices for both final models will be discussed in the following section.

Final path models. The final composite and the final combined models, shown in Figures 8 and 9, support the role that GCA and executive function skills play in partially mediating the relation between AGE and arithmetic performance. Most paths leading to

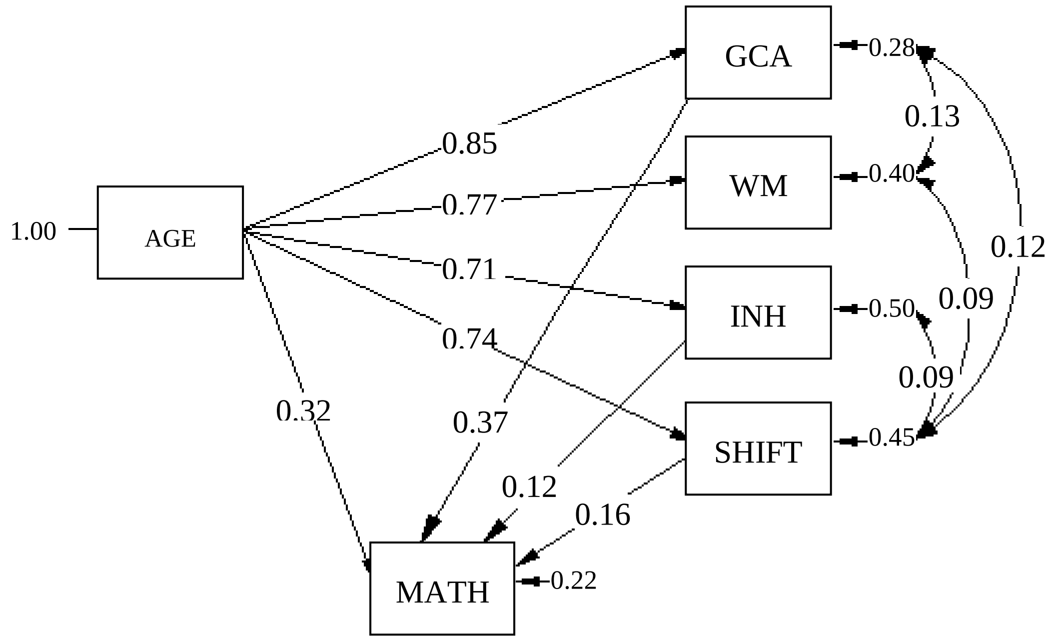


Figure 8. Final composite model with standardized coefficients and residuals.

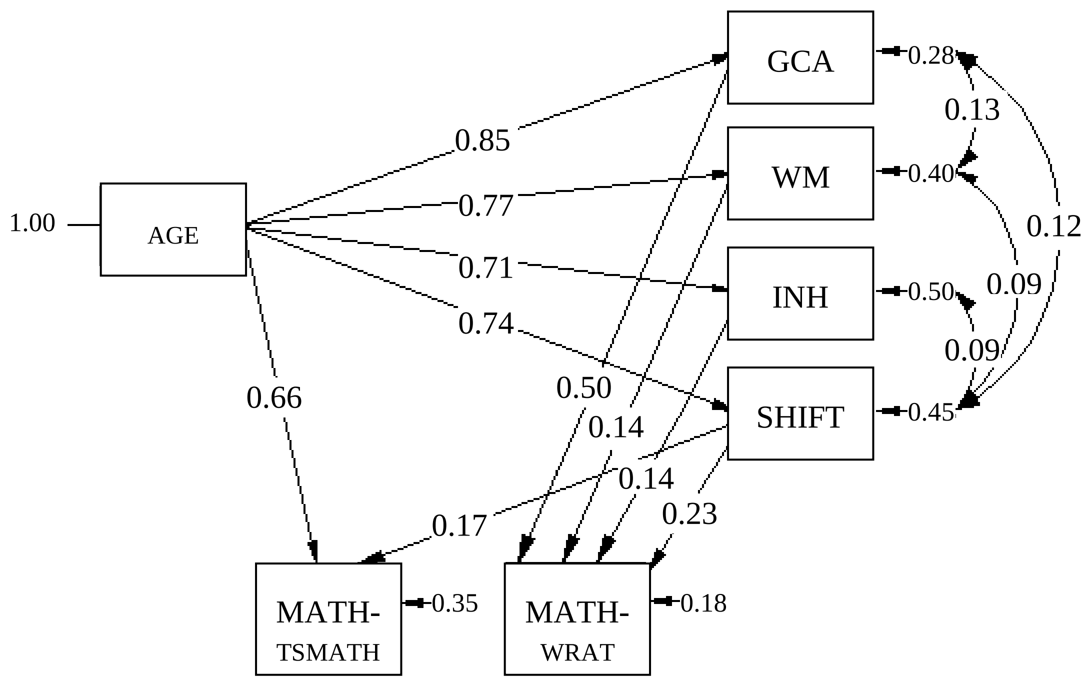


Figure 9. Final combined model with standardized coefficients and residuals.

MATH in the composite model were found to significantly improve the fit of the model and AGE, GCA, WM, INH, and SHIFT each contributed in some way to the prediction of MATH-WRAT and MATH-TSMATH in the combined model. The fit statistics associated with the final models indicate that the final composite model was a better fit to the data than the final combined model (see Table 21). However, the data used to form the correlation matrices for each analysis differed slightly as the composite matrix included MATH and the combined model included MATH-WRAT and MATH-TSMATH; thus, the composite model cannot be compared to the combined model as if it were a nested version because the correlation matrices differed across models.

In terms of comparing fit indices between the two models, the significance of the χ^2 value associated with the final combined model indicates that there may have been problems with the fit of this model. Specifically, the *p*-value associated with the χ^2 of the final composite model was non-significant while the *p*-value associated with the χ^2 of the final combined model was significant at $p < .001$. Kline (2005) noted that the *p*-value of the χ^2 of a model often is not a good measure of fit as χ^2 values tend to be inflated with larger models. Even though the χ^2 of the combined model was significant, it does not necessarily indicate poor fit. However, the RMSEA values associated with both final models indicate that the composite model was, in fact, a better fitted model than the combined model, as the RMSEA for the composite was less than .05 and the RMSEA for the combined was greater than .10. The comparative fit index (CFI) and the standard root mean square residual (SRMR) values were within acceptable range for both of the final models. Thus, though the results from both models will be reviewed in the coming paragraphs, the results from the combined model should be interpreted with caution.

The standardized path and correlation coefficients generated in the final composite model, represented here by their respective *z*-scores, indicate that AGE was a better predictor of GCA, WM, INH, and SHIFT ($z = 19.58$, $p < .001$, $z = 14.77$, $p < .001$, $z = 12.18$, $p < .001$, $z = 13.28$, $p < .001$, respectively) than it was to MATH ($z = 3.68$, $p < .001$). The smaller (though still significant) link between AGE and MATH in the composite model is accounted for by the finding that AGE was not significantly predictive of MATH-WRAT performance but, in fact, was the strongest predictor of MATH-TSMATH performance in the combined model ($z = 8.99$, $p < .001$). Thus, the results from the combined model imply that age does not predict all aspects of arithmetic performance; age may only predict arithmetic fluency and not proficiency. These results are consistent with the regression analyses and the implications will be discussed in the discussion section.

As both final models indicted that the link from AGE to arithmetic was partially or fully mediated by executive function skills and/or general cognitive abilities, the effects of AGE on arithmetic were both mediated (a form of indirect) and direct (see Holmbeck, 1997). In the final composite model, the direct effect of AGE on MATH was significant ($\gamma = 0.31$, $z = 3.68$, $p < .001$), and the Sobel *z* test (Sobel, 1982; 1988) was significant for the meditational paths linking AGE to MATH through GCA ($z = 4.52$, $p < .001$), INH ($z = 1.97$, $p = .05$), and SHIFT ($z = 2.61$, $p = .009$), suggesting that all three of these variables partially mediated the relation between AGE and MATH. In sum, the total effects of AGE on MATH were reduced by 13.3% after accounting for mediating effects. In the final combined model, the direct effect of AGE on MATH-TSMATH was significant ($\gamma = 0.66$, $z = 8.99$, $p < .001$), though the direct effect of AGE on MATH-WRAT was not. The Sobel *z* test was significant for the path linking AGE to MATH-TSMATH through SHIFT ($z = 2.38$, $p = 0.02$), suggesting that SHIFT partially mediated the relation between AGE and MATH-TSMATH. In this model, the total effects of AGE on MATH-TSMATH were reduced by 4.4% after accounting for the mediating effects of SHIFT. In addition, the Sobel *z* test revealed that each indirect path linking AGE to MATH-WRAT through GCA ($z = 6.77$, $p < .001$), WM ($z = 2.31$, $p = .02$), INH ($z = 2.72$, $p = .006$), and SHIFT ($z = 3.66$, $p < .001$), was significant, indicating that these variables fully mediated the relation between AGE and MATH-WRAT.

SHIFT was the only executive function variable to significantly predict MATH-TSMATH performance in the combined model ($z = 2.34$, $p = .02$), and was the strongest executive function predictor of both MATH scores in the composite model ($z = 2.59$, $p = .01$) and MATH-WRAT performance in the combined model ($z = 4.48$, $p < .001$). While INH and WM both were significantly predictive of MATH-WRAT scores in the combined model ($z = 2.92$, $p < .004$, $z = 2.81$, $p = .005$, respectively), only INH remained significantly predictive of MATH in the composite model ($z = 2.17$, $p = .03$). In addition, although GCA was the strongest predictor of MATH-WRAT scores in the combined model ($z = 9.79$, $p < .001$) and of MATH scores in the composite model ($z = 4.61$, $p < .001$), it was not a significant predictor of MATH-TSMATH scores in the combined model.

As the pattern of correlations was consistent between models, the correlation coefficients from both models will be treated as interchangeable. The strongest correlations between predictor variables were between GCA and WM ($z = 4.50$, $p < .001$) and between GCA and SHIFT ($z = 4.04$, $p < .001$). While WM and INH were both significantly correlated with SHIFT ($z = 2.56$, $p = .01$, $z = 2.40$, $p = .02$, respectively), INH was not significantly correlated with any other predictor variable.

Discussion

This study examined the influences of age, general cognitive ability, and executive functions on different forms of arithmetic performance across a range of individuals in middle childhood, late childhood, early adolescence and young adulthood. The hypotheses were that: 1) age, general cognitive ability, and executive functions (working memory updating, inhibition, and set shifting) would directly predict arithmetic scores; 2) age would predict differences in general cognitive ability and executive functions; 3) executive functions would continue to significantly predict arithmetic scores after accounting for the effects of age and general cognitive ability, and; 4) each of these components (e.g. age, general cognitive ability, and executive functions) would add to the overall prediction of arithmetic scores. All of the hypotheses examined in this study were at least partially supported by the results from regression and path analyses.

Age, General Cognitive Ability, and Arithmetic

The hypothesized positive linear relation between arithmetic performance and age and general cognitive ability (Hypothesis 1) was fully supported by results from the correlation analyses but only partially supported by the results from the regression and path analyses. In terms of the regression and path analyses that included MATH as the outcome variable, age was found to significantly contribute to MATH even after accounting for the influences of both general cognitive ability and executive functions, though these processes were significant partial mediators. While the same pattern of results held for the significant predictive relation between MATH and general cognitive ability in the path analysis, when the variables that made up the construct of general cognitive ability were parceled out in the regression model, GCA-RAVENS was the only measure of this construct to significantly account for overall arithmetic performance, i.e., MATH. These results indicate that non-verbal aspects of fluid intelligence, as measured by Raven's Standard Progressive Matrices, may be more closely linked to one's overall arithmetic skills than verbal aspects, as measured by the GCA-BPVS.

The results from the analyses conducted with each measure of arithmetic performance separately (rather than averaged together) indicated that age and general cognitive skills related to arithmetic performance in different ways. In the regression analyses, AGE significantly contributed to MATH-TSMATH performance when the variance from general cognitive skills and executive functions had been accounted for; however, in the MATH-WRAT regression models, AGE did not significantly contribute to arithmetic proficiency when entered with variables related to general cognitive skills nor when entered with executive function variables. In addition, GCA-RAVENS and GCA-BPVS scores were found to significantly contribute to MATH-WRAT scores but not MATH-TSMATH scores. In fact, in reference to the semi-partial correlations, the unique contributions of GCA-BPVS to MATH-TSMATH scores were negative, while the unique contributions of GCA-BPVS to MATH-WRAT were positive. Thus, combining MATH-WRAT and MATH-TSMATH into a composite variable neutralized the contributions of GCA-BPVS to arithmetic. These results are consistent with the results from the path analyses that analyzed general cognitive ability as a composite variable, i.e., GCA rather than as two separate variables, i.e., GCA-BPVS and GCA-RAVENS. The inconsistent predictability of age and general cognitive skills to different outcome measures of arithmetic may be due to fundamental differences in what each measure of arithmetic actually measured.

Arguably, one's arithmetic skills may be not only a reflection of the accuracy, or proficiency, with which one can perform mathematical computations, but the relative speed, or fluency, with which one can continue to perform these actions while maintaining a consistent level of accuracy (e.g., Geary et al., 1991; Kaye et al., 1989; Ramos-Christian et al., 2008). As proficiency is often denoted by one's final score on a mathematical exercise, it stands to reason that proficiency represents the most basic measure of achievement in mathematics. As the MATH-WRAT task measures one's proficiency in performing arithmetic computations, i.e., addition, subtraction, multiplication, logarithms, etc., the MATH-WRAT may be viewed as a measure of arithmetic proficiency. The TSMath task, on the other hand, not only measures one's proficiency in performing simple computations, but also measures the amount of time it takes to perform these actions. Thus, TSMath allows for one's arithmetic fluency to be computed from the speed at which one is able to perform mathematical computations accurately. Based on the results from the combined path analyses and the separate MATH-TSMATH and MATH-WRAT regression analyses, it seems that greater general cognitive skills contributed to greater achievements in arithmetic regardless of age. Moreover, with age, individuals were able to perform simple mathematical computations more efficiently while maintaining a consistent degree of accuracy. These results are supported by previous findings that have reported a positive linear relation between age and general processing speed among children and young adults (Kail, 1991; specifically pertaining to arithmetic fluency: Ashcraft, 1982; Geary et al., 1991; Kaye et al., 1989), and tie in to studies that have reported a similar linear trend between age and executive function efficiency (Cepeda et al., 2001; Ellefson et al., 2006; Reimers & Maylor, 2005). Also of importance is that while TSMath measures fluency in computing simple addition and subtraction, the WRAT may be defined as a measure of the extent of one's knowledge in arithmetic, as it contains both simple and complex arithmetic problems. In the future it would be helpful to compare MATH-WRAT performance with performance on a measure of arithmetic fluency that contains complex computations, as well, to determine to degree to which the different results found here may be explained by fluency versus proficiency differences rather than simple versus complex task differences.

Age and Executive Functions

As prior studies have found that executive functions undergo age-related changes (Cepeda et al., 2001; Frye et al., 1995; Huizinga & van der Molen, 2007; Kray et al., 2004), the link between age and executive functions may be important to understanding the link between executive functions and arithmetic, as the latter may be influenced by changes in the former. In the current study, the link between age, working memory updating, inhibition, and set shifting was examined through correlation and path analyses. The significant main effects of age for all variables and the positive non-parametric correlations found between age and each variable related to executive function indicated that, in general, older participants performed significantly better on all measures of executive functions than younger participants, supporting Hypotheses 1 and 2. ANOVAs comparing age group differences along with their subsequent post-hoc tests indicated that performance on all measures followed a general linear pattern such that participants in the older age groups performed better than participants in the younger age groups. In addition, the results from the path analyses indicated that age significantly contributed to differences in general cognitive ability, working memory updating, inhibition, and set shifting such that as the age of the participants increased, scores in each of these constructs increased as well. These results support previous evidence that components of working memory updating, inhibition, and set shifting display signs of age-related improvements across middle childhood and often continue to improve until late adolescence or early adulthood (Huizinga et al., 2006; Luciana et al., 2005; Reimers & Maylor, 2005). In relation to arithmetic performance, the results from this study indicate that age-related changes in general cognitive ability and executive functions add to the link between age and arithmetic.

Executive Functions and Arithmetic

As previous studies have found a link between executive functions and arithmetic skills in young children (Bull et al., 2008; Bull & Scerif, 2001; Espy et al., 2004; Mazzocco & Kover, 2007; St Clair-Thompson & Gathercole, 2006; van der Sluis et al., 2007), the primary intention of this study was to examine the relation between arithmetic skills and executive functions in a sample of individuals representing middle childhood, late childhood, early adolescence, and early adulthood. Thus, it was hypothesized that greater executive function skills in the areas of working memory updating, inhibition, and set shifting would predict greater arithmetic skills in a sample of older children and adults, as well (Hypotheses 1, 3, and 4). The results from the analyses of non-parametric correlations between each measure of executive function and each measure of arithmetic, and between each executive function composite variable and the arithmetic composite, provided full support for Hypothesis 1 on both the indicator, i.e., the separate variables, and the construct, i.e., the composite variables, levels. These results indicated that, in general, working memory updating, inhibition, and set shifting skills were positively associated with arithmetic skills within this sample of individuals.

Also, as individual differences in executive function skills exist apart from differences in general cognitive skills (Blair, 2006) and across individuals of the same age (Scope, Empson, & McHale, 2010), it was hypothesized that greater executive function skills would predict greater arithmetic skills regardless of one's age or general cognitive ability (Hypothesis 2). The results from the regression analyses offered partial support for this hypothesis; after controlling for age and general cognitive ability, components of working memory updating and set shifting contributed to arithmetic proficiency, i.e., MATH-WRAT scores. Specifically, increases in scores on the Tic-Tac-Toe task, but not the Recall of Digits task, and increases in Shape School Switch scores, but not Figure Matching, contributed to increases in MATH-WRAT scores. This difference could be due to the Tic-Tac-Toe task and the Recall of Digits task tapping different components of the working memory system. The visually-presented Tic-Tac-Toe task likely involves the visuo-spatial sketchpad – responsible for recognizing and temporarily storing visual information – while the orally-presented Recall of Digits task may rely on the phonological loop – responsible for buffering auditory verbal information (Baddeley & Hitch, 1974). These results are consistent with the results from Bull et al. (2008), which found that preschoolers' visuo-spatial working memory skills significantly predicted math achievement in kindergarten, first grade, and second grade. Thus, it may be that among older children and adults, as well, visuo-spatial – rather than phonological – working memory processes contribute to arithmetic proficiency or achievement.

As the Shape School Switch and Figure Matching tasks were quite similar in terms of task demands, differences in task administration and format may have accounted, in part, for the finding that Shape School Switch contributed to arithmetic proficiency and Figure Matching did not. As both the WRAT and the Shape School tasks were paper-based and administered by an experimenter and the Figure Matching task was computer-based and self-administered, subtle relations between set shifting and arithmetic proficiency may have been masked by significant differences in measurement characteristics (between the Figure Matching and WRAT tasks). In addition, the Shape School required verbal responses while Figure Matching required manual responses, the Figure Matching task presented participants with four items on the screen at a time (one trial at a time) while the Shape School presented participants with all 48 trials of each block at one time, and participants' RTs on the Figure Matching task were recorded on every trial while RTs on the Shape School task were recorded by block. In any case, these results support previous studies that have found links between components of working memory and performance in arithmetic (Bull et al., 2008; Bull & Scerif, 2001; Gathercole & Pickering, 2000; St-Clair-Thompson & Gathercole, 2006) and lend further credence to the lesser-reported finding that set shifting skills influence arithmetic achievement, specifically in terms of proficiency (Bull & Scerif, 2001) – at least in a sample of older children and adults.

While components of working memory updating and set shifting were found to significantly contribute to arithmetic proficiency after controlling for age and general cognitive ability, arithmetic fluency – represented by RTs to correct trials on the TSMath task (MATH-TSMATH) – was influenced by scores on one a measure of inhibition (Shape School Inhibition) and not by measures of working memory updating or set shifting. One explanation may be that efficiency on the inhibition block of the Shape School accounted for significant variance in arithmetic fluency because the variables taken from both the Shape School and the TSMath tasks reflect components of cognitive efficiency or general processing speed; in other words, shared RT and accuracy demands between the two tasks may have accounted for their shared variance. However, if cognitive efficiency or processing speed demands accounted for these results then one would expect to find that all measures of

RT or efficiency would significantly contribute to MATH-TSMATH, as well. In fact, RT and efficiency measures from several other tasks (Tic-Tac-Toe, Stop-Signal, Shape School Switch, and Figure Matching) did not significantly contribute to MATH-TSMATH. Thus, the fact that both of the measures – INH-SS-INH and MATH-TSMATH – were measures of RT and accuracy simply may have reduced the amount of extraneous measurement variance that would have clouded the relation between inhibition and arithmetic and allowed for the relation to emerge significant in this model even though it was not significant in the MATH-WRAT model. These results lend partial support to the documented link between inhibition and arithmetic skills among younger children (e.g., Bull & Scerif, 2001; Espy et al., 2004) and expand this link to older children and adults.

While the results from the path analysis involving the composite MATH variable were highly consistent with the results from the regression analysis with the same DV, the final path model that included both arithmetic variables differed in several ways from the results of the MATH-WRAT and MATH-TSMATH regression analyses. For example, though MATH-WRAT performance was significantly influenced by general cognitive ability (both GCA-RAVENS and GCA-BPVS), working memory updating (WM-TTT), and set shifting (SS-SS-SW) in the regression model, it was influenced by inhibition (INH) in the final combined path model. Thus, when performance was averaged across both the Shape School Inhibition task and the Stop-Signal task, a significant relation emerged between inhibition and arithmetic performance, lending further support to previous findings that greater inhibition skills relate to greater arithmetic skills (Bull & Scerif, 2001; Espy et al., 2004; St. Clair-Thompson & Gathercole, 2006).

The relation between MATH-TSMATH and executive functions was not consistent across the regression and path models. In the MATH-TSMATH regression model, arithmetic fluency was significantly influenced by inhibition (INH-SS-INH); yet, in the combined path model, fluency was influenced by set shifting (SHIFT) and not by inhibition (INH). It may be that when inhibition performance was averaged across both measures of inhibition, the significant association between inhibition efficiency and arithmetic fluency was weakened by the presence of variance from a non-efficiency/fluency based measure (INH-SSRT). However, when both efficiency measures of set shifting (SS-SS-SW and SS-FIGURES) were averaged, a significant relation emerged between set shifting composite scores and MATH-TSMATH scores that reflected the association between the efficiency measures of set shifting and the fluency measure of arithmetic. Thus, while the significant relation between INH-SS-INH and MATH-TSMATH in the regression analysis may have represented more than simply shared efficiency/fluency demands, the significant path from set shifting to MATH-TSMATH in the final combined path model might have resulted from the shared task demands.

As previous studies that have examined associations between the latent constructs of executive function have found evidence suggesting that these constructs are separate but related (Anderson et al., 2001; Huizinga et al., 2006; Miyake et al., 2000; St. Clair-Thompson & Gathercole, 2006), it was hypothesized that the final path model would include correlations between each of the executive function composite variables. This hypothesis was supported by the results, with the exception of the non-significant correlation between WM and INH. In addition, the final models indicated that GCA was significantly correlated with WM and SHIFT. These results are partially consistent with previous literature that has reported correlations among executive function variables (e.g., Miyake et al., 2000) and between executive function variables and general cognitive ability (e.g., Friedman et al., 2006); however, the finding that the construct of inhibition was not significantly correlated with the construct of working memory updating does not support previous results (Miyake et al., 2000). The results imply that although executive functions largely are distinct from each other and from general intelligence, they are not entirely independent cognitive processes.

Relation to previous literature. Across studies, there have been inconsistencies in reports of a link between executive functions and arithmetic (e.g., Bull et al., 2008; Gathercole & Pickering, 2000; Tolar et al., 2009). The results from this study indicate that the link between executive functions and arithmetic varies depending on the executive function component that is targeted and on the way arithmetic is measured. In general, as different aspects of executive function were found to relate to arithmetic performance in different ways, and as participants displayed differential patterns of performance across the executive function measures, the results from this study seem to support the diversity – rather than unitary – perspective of executive function (e.g., Anderson et al., 2001; Huizinga et al., 2006; Huizinga & van der Molen, 2007; Miyake et al., 2000; Welsh et al., 1991). In addition, the significant correlations between most of the executive function composite variables in the path analyses lend support to the idea that, although executive functions are largely distinct from each other, they also may overlap (see Miyake et al., 2000).

Working memory. Similar to the significant link between working memory updating and MATH-WRAT performance found in the current study, in a study that tested children on multiple measures of working memory and standardized measures of vocabulary, literacy, and arithmetic, Gathercole and Pickering (2000) also reported a link between processes related to working memory updating and arithmetic achievement in children at seven years of age and again at eight years of age. That the findings from the current study are consistent with the results from Gathercole and Pickering's (2000) study may relate to similarities in the measurement of arithmetic proficiency or achievement across both studies. In the current study and in Gathercole and Pickering (2000), arithmetic proficiency was measured by accuracy (raw scores) on a standardized measure of basic arithmetic skills. In addition, St. Clair-Thompson and Gathercole (2006) also reported a link between performance on a standardized achievement measure of mathematics and performance on tasks of working memory among 11-year-olds, indicating that accuracy or proficiency in arithmetic is related to working memory skills among older children and further supporting the results of this study.

Unlike the results from the current study, Tolar et al. (2009) found that working memory processes were significantly related to computational fluency skills among a sample of college students and that computational fluency mediated the relation between working memory and algebra achievement. In the current study, working memory updating skills did not significantly predict arithmetic fluency. The different pattern of results found in the current study versus the results found by Tolar et al. (2009) may be due to several factors. Most notably, the participants in Tolar et al.'s sample all were between the ages of 18 and 25, i.e., young adults, while the participants in the current sample ranged from 6 to 39 years of age. Perhaps, the relation between working memory and computational fluency reported by Tolar et al. (2009) is most pronounced during young adulthood, and young adults made up less than one-quarter of the sample in the current study.

Although the current study only examined two tasks of working memory updating, Tolar et al (2009) used four updating tasks: reading span, counting span, backwards digit recall, and letter-number sequencing. Had the current study included more measures of working memory updating, perhaps the facets of updating that may be related to arithmetic fluency would have been illuminated more clearly. Finally, fluency was not measured consistently between studies, i.e., between the current study and Tolar et al. As mentioned above, the fluency measure in the current study only measured fluency for simple, single-digit addition and subtraction problems. In Tolar et al., several measures – including multi-digit computations, addition, subtraction, multiplication, and division – were used to tap the construct of fluency. Thus, future attempts to examine the link between working memory updating and arithmetic fluency should not only include a large developmental sample, but also both simple and complex measures of fluency. Such efforts would inform the question of whether or not the results from this study – that working memory updating skills relate to arithmetic proficiency and not fluency – are consistent across different contexts and age groups.

Inhibition. Along similar lines, differences in the results from this study regarding the link between inhibition and arithmetic and the results from previous studies that have not found a consistent link between inhibition and arithmetic performance among older children (e.g., Bull et al., 2008; Van der Sluis et al., 2007) may be due to differences in the way arithmetic and inhibition were measured across studies. Though, in the current study, a significant link was found between inhibition and fluency – but not proficiency – in arithmetic, neither Bull et al. (2008) nor van der Sluis et al. (2007) included a measure of arithmetic fluency in their research design. In both of these studies, arithmetic was measured by accuracy scores on standardized measures of arithmetic achievement (a measure of proficiency). Considering that both previous studies did not find that inhibition skills predicted arithmetic achievement (proficiency) among older children, the results from these two previous studies are consistent with the results from the current study; though the previous studies did not contain measures of arithmetic fluency and, therefore, cannot be compared exactly to the current study. Thus, adding to the statement above concerning future studies on working memory and arithmetic, future studies examining the link between inhibition and arithmetic fluency among older children, adolescents, and adults are needed in order to explore the reliability of the results from this study.

In contrast to the studies presented above, several other developmental studies have found that inhibition skills significantly predict arithmetic proficiency from preschool to late childhood (e.g., Bull & Scerif, 2001; Espy et al., 2004; St.-Clair-Thompson & Gathercole, 2006). For example, Bull and Scerif (2001) administered a standardized measure of single and multi-digit computational proficiency to preschoolers and St. Clair-Thompson and Gathercole (2006) used the national attainment test scores of 11-year-olds as a measure of arithmetic achievement. In both studies, arithmetic proficiency was predicted by inhibition skills. However, unlike the current study, Bull and Scerif used only the Stroop task to measure inhibition and may not have captured the same aspects of inhibition that were captured in the current study. Though St. Clair-Thompson and Gathercole (2006) used both the Stop-Signal and Stroop tasks to measure inhibition, they included only one age group in their research design, in contrast to the current study that included four age groups. Therefore, the significant relation found between arithmetic proficiency and inhibition by St. Clair-Thompson and Gathercole may be prominent among 11-year-olds but not among other age groups, i.e., the younger and older age groups represented in the current study. In the future, the reliability of the results from this study pertaining to the non-significant relation between inhibition and arithmetic proficiency could be examined by conducting a similar study with a developmentally diverse sample using a comprehensive and valid battery of inhibition tasks and an accuracy-based measure of arithmetic achievement.

Set Shifting. Results from this study pertaining to the link between set shifting and arithmetic proficiency were not consistent with the previous literature on all accounts; however, previous literature has been mixed in terms of the relation between the two constructs, and no studies to date have examined the relation between set shifting and arithmetic fluency. The results from the current study indicated that stronger set shifting skills predicted greater proficiency in arithmetic, even while controlling for age and general cognitive ability. The present findings are inconsistent with previous studies that have not reported a significant link between set shifting and arithmetic achievement or proficiency among younger children (e.g., Espy et al., 2004), children in middle childhood (four-to-eight-years; Bull et al., 2008), or children in late childhood (e.g., Lee et al., 2009; St.-Clair-Thompson & Gathercole, 2006). Although these previous studies did not include adults in their samples (unlike the current study), an examination of the pattern of correlations – separated by age group – between set shifting measures and MATH-WRAT scores reveals that aspects of set shifting were significantly related to arithmetic proficiency across all age groups in the current study (see Table 11). Thus, the inclusion of adults in the current study cannot, alone, account for the differences between previous findings and the present findings.

A possible explanation for the inconsistencies between the current findings and the results from previous studies may be found in the way that arithmetic proficiency was measured in this study. While Espy et al. (2004) did not find a significant relation between set shifting and arithmetic proficiency among preschool children and Bull et al. (2008) also found this relation to be non-significant among four-to-eight-year-olds, Bull et al. suggested that these findings may have been due to the simplicity of the arithmetic measures that were included in both studies. Though Espy et al. (2004) and Bull et al. (2008) used standardized measures of arithmetic that assessed counting, number recognition, addition, and subtraction, other studies that have used more complex measures of arithmetic, i.e., those that included single and multi-digit addition problems, subtraction, multiplication, and division, have indicated that set shifting is significantly related to arithmetic proficiency among four-to-five-year-olds (van der Sluis et al., 2004) and seven-to-nine-year-olds (Bull & Scerif, 2001; McLean & Hitch, 1999). As the current study measured arithmetic proficiency using a standardized measure that included both simple and complex arithmetic problems, the finding that set shifting is related to arithmetic proficiency adds support to the theory that set shifting skills are more highly related to complex – rather than simple – forms of arithmetic.

Other inconsistencies in the results from the current study and the results from previous studies (e.g., Lee et al., 2009; St.-Clair-Thompson & Gathercole, 2006) may be linked to differences in methodologies, as well. Lee et al. (2009) used two tasks to measure set shifting (number-letter and plus-minus tasks; adopted from Miyake et al., 2000) but, in contrast to the current study, did not measure performance by combining measures of accuracy and RT. Instead, Lee et al. used participants' average switch costs across the two tasks as measures of set shifting, i.e., the difference between average switch trial RT and average non-switch trial RT. In addition, Lee et al. included only one age group in their sample (11-year-olds) while the current study included four. Moreover, in a study that also involved 11-year-olds and utilized the letter-number and plus-minus tasks, St.-Clair-Thompson and Gathercole (2006) conducted a preliminary principle components analysis to determine the factor structure of all of their executive function variables and found that both set shifting tasks failed to load onto any of the factors; thus, they decided not to include these variables in their subsequent analyses. As a result, this study did not directly examine the relation between set shifting and arithmetic. Therefore, although Lee et al. (2009) and St.-Clair-Thompson and Gathercole (2006) reported that set shifting was not significantly related to proficiency in complex arithmetic problem solving while the current study found that set shifting was significantly related to arithmetic proficiency, these inconsistent results may be due to methodological differences across these three studies.

In summary, the majority of the results from this study support previous findings. Age group differences and inter-measure performance differences (within individuals) found here support previous accounts of the diversity of executive functions (e.g., Miyake et al., 2000; Welsh et al., 1991). The present finding that working memory updating skills predict proficiency in arithmetic across middle and later childhood is supported by the previous findings of Gathercole and Pickering (2000) and of St. Clair-Thompson and Gathercole (2006), and extend these findings to early adolescents and adults. Like Bull and Scerif (2001) and St. Clair-Thompson and Gathercole (2006), the current study found that inhibition skills predict arithmetic skills; however, unlike previous findings, inhibition skills predicted only fluency and not proficiency in arithmetic. As the literature is mixed regarding the role of set shifting in arithmetic, with some studies reporting a link among children in early and middle childhood (Bull & Scerif, 2001; McLean & Hitch, 1999; van der Sluis et al., 2004) and others reporting no such link among older children (Lee et al., 2009; St.-Clair-Thompson & Gathercole, 2006), results from this study – that set shifting predicted proficiency in arithmetic – were partially consistent with the previous literature. Overall, the results from this study were difficult to compare against results from previous studies given the lack of prior evidence regarding the link between executive functions and arithmetic fluency.

Limitations to the Study

Common to all psychological research, there were various limitations to this study, the first being that this study was based on data that had been previously collected. As such, the measures were not chosen specifically with this study in mind. Problems existed within and across certain measures that prevented the data from being analyzed using more complex latent variable analyses, i.e., structural equation modeling (SEM). Prior to conducting path analyses, models using SEM were attempted in which age represented a latent variable with a single predictor and working memory updating, inhibition, set shifting, general cognitive ability, and arithmetic represented five additional latent variables each with two indicators. Due to problems with high multicollinearity, the results from these analyses were invalid. Combining the indicator variables to form composite variables alleviated the multicollinearity problem; thus, path analysis was preferred over SEM for this dataset. The disadvantage of using path analysis over SEM is that specific variance within an indicator can be masked once it is combined with the variance from another indicator. For example, if an individual were to perform highly above average on one measure and highly below average on another measure, the composite score for these two measures would indicate that this individual demonstrated average performance on this particular construct when, in fact, the individual was well beyond the average range on either of the two measures within the construct. Additional measures of each latent construct might have allowed for certain highly correlated variables to be dropped from the SEM analyses while still allowing for each latent construct to correspond to at least two indicators (apart from age), thus making it possible to use SEM rather than path analysis.

In addition, sample size limitations did not allow for separate hierarchical regressions or path models to be analyzed for each age group. As there were a limited number of participants in each age group (fewer than 50 per group), there was not enough power to conduct multi-sample path (or SEM) analyses. However, when correlations between arithmetic variables and experimental variables were separated by age group, a number of age-group differences emerged. For example, though the relation between WM-TTT and MATH-WRAT remained consistent across age groups, the relation between WM-DIGITS and MATH-WRAT was significant among younger children but not older children or adults, indicating that some aspects of working memory updating might be more highly related to arithmetic proficiency in younger age groups than in older age groups. Thus, a multi-sample analysis could have revealed potentially significant differences in the way each form of cognition affected arithmetic skills according to age group, i.e., the moderating effects of age (see Holmbeck, 1997).

Although this study included a wide range of age groups, the sample essentially skipped-over an important period of development, namely, middle adolescence. As studies have shown that executive function skills develop throughout adolescence (Huizinga et al., 2006; Huizinga & van der Molen, 2007; Luciana et al., 2005), expanding the sample even further to include a group of mid-adolescents may have generated a more comprehensive and complex picture of age-related variance. For this reason, this study not only would have benefited from including more participants in each age group (at least 100 per group), but from including at least one additional age group representing mid-adolescence, i.e., between 15 and 16 years of age, as well.

Because this was a cross-sectional study, even if the sample had been large enough to examine predictability patterns separately for each age group, such analyses would not have reflected differences due to developmental change, as different participants were in different age groups. The cross-sectional nature of this sample also could have led to cohort effects, as the age effects observed in this study might be different for different groups of individuals through time. Though conducting this study longitudinally would reduce these limitations, a longitudinal study of this nature would be subject to issues of attrition, additional time and cost burdens, and testing effects. Consequently, conducting this study longitudinally would have eliminated many restrictions caused by the cross-sectional sampling method but would have resulted in other limitations that also could have compromised the validity of the study.

In regards to the generalizability of the data, another possible limitation is that the results from this study may not generalize to different settings or environments. The data were collected in an isolated environment by experimenters who followed strict standard procedures. In the school setting, arithmetic skills are assessed by a number of measures, i.e., non-standardized tests, individual and group assignments, take-home assignments, projects, etc. In this study, arithmetic skills were measured by a standardized achievement test and by a computer-based experimental task switching task involving simple addition and subtraction. Thus, scores on the measures of arithmetic that were included in this study may not represent accurate reflections of the arithmetic skills that children and adults may demonstrate in real-world environments. If this study had examined scores from school-based measures of arithmetic in addition to standardized and experimental measures, the results may have provided more information about the nature of these processes as they occur in the real-world.

A final limitation to this study is in the way that the data were prepared for analysis.

Though the variables in the current study were transformed on an as-needed basis, in a recent statistics text, Field (2009) suggested that if one variable in a dataset requires transformation, all variables in the set also should be transformed. In addition, he stated that all variables should be transformed using the same transformation method. He argued that transforming one variable in a dataset and not the others may not affect the relationships between variables but could affect the specific differences between the variables, as they would no longer be of the same measurement units; thus, transforming all the variables in a dataset in the same way allows all variables to be in the same units. Therefore, the results from this study may have been affected by the fact that only a portion of the variables were transformed and two different transformations were used.

Implications and Directions for the Future

Numerous factors contribute to the learning and practice of arithmetic across childhood and early adulthood. Although a strong foundation in arithmetic and knowledge of mathematical concepts has been found to contribute to success beyond the school years (Paglin & Rufolo, 1990; Rivera-Batiz, 1992), the neurocognitive and developmental processes that influence arithmetic performance remain somewhat unclear. The relative ability and efficiency with which one may learn and carry-out functions of arithmetic may be influenced by general processes, such as age and general cognitive abilities, but also by the functioning and development of higher-order cognitive processes such as executive functions (Espy et al., 2004; Fuchs et al., 2006; Griffin et al., 1994; Jordan et al., 1995; Kroesbergen et al., 2009). Moreover, different types of mathematical situations requiring solution formation may depend on different executive functions, and the role of executive functions in arithmetic performance may further depend on age. However, the extent to which age moderates differential relations between executive functions and arithmetic performance remains an important future area of study.

The results from this and previous studies that have found a link between executive functions and arithmetic achievement may have important implications for individuals with significant learning difficulties, specifically in the area of arithmetic. An implication of the findings from the current study is that children and young adults who are less skilled in the areas of working memory, inhibition, and set shifting may find it more difficult to mentally navigate through strings of digits and computations than others who are more skilled in these areas of executive function. As such, it is not surprising that children with arithmetical difficulties have shown similar difficulties in areas of executive function (Bull, Johnston, & Roy, 1999; Bull & Scerif, 2001; McLean & Hitch, 1999). By examining this topic in greater detail, through longitudinal methods and/or with larger and more comprehensive cross-sectional samples, the fields of education and psychology may gain a more complete understanding of the sources and development of arithmetic difficulties that typically manifest during the school years.

Though not explicitly addressed by the current study, the neurological developments that are associated with age-related improvements in arithmetic skill and executive functions represent another important area for future study. The neurological systems related to both arithmetical processes and executive functions tend to display similar patterns of age-related shifts from general to specific and from global to local areas of activity (e.g., Crone et al., 2006b; Durston et al., 2002; Rivera et al., 2005; Tamm et al., 2002). Hence, studying the similarities and differences in the patterns of neural development that accompany cognitive developments in the areas of arithmetic and executive functions may provide further insight into the link between executive functions and arithmetic. Although the relation between both areas of neurocognitive development may be inferred by reviewing the separate branches of literature, it is difficult – if not impossible – to find a comprehensive developmental study that examines the neurological underpinnings of skill development in both arithmetic and executive functions. Moreover, no known studies to date have examined neurological differences in the development of arithmetic proficiency versus arithmetic fluency. Future studies of this nature would extend our knowledge of the typical and atypical patterns of development in both areas of arithmetic and would inform what is known about the relation between developments in arithmetic skill and developments in executive functions.

In sum, results from this study indicate that dynamic components of cognition, i.e., verbal and non-verbal components of intelligence and multiple components of executive function, influence arithmetic performance across middle and late childhood, early adolescence, and early adulthood, but that this influence is not consistent across different contexts. Poor working memory updating, inhibition, and/or set shifting skills – both independently and collectively – can result in poor arithmetic performance regardless of one's age or general cognitive skills. However, as evidenced by the findings from the current study, executive functions are not universally applied in all arithmetic tasks and across all age groups; thus, predictions about how executive functions influence arithmetic achievement should be specific to age and context. Unfortunately, in our current state of understanding, we know almost nothing about how executive functions predict arithmetic fluency, specifically, even though fluency is often tapped in the school setting by way of timed tests. Therefore, more evidence of the individual and shared roles of executive functions in arithmetic proficiency and fluency across age groups is needed.

Executive function skills typically improve with age, evidenced by the fact that the older individuals in the sample demonstrated better executive function skills than the younger. As executive functions are malleable, dynamic components of cognition that undergo age-related change, they may be improved through targeted education efforts, as well. At this point, more studies are needed in order to identify how executive functions differentially relate to arithmetic proficiency and fluency at different ages. In general, this study suggests that developing a better understanding of the relation between executive functions and school achievement may provide insight into the cognitive mechanisms that allow learning to occur in the classroom and could open the possibility that an executive function-inspired curricula might help improve achievement in school (e.g. Tools of the Mind program, Bodrova & Leong, 1996, see Diamond, Barnett, Thomas, & Munro, 2007).

List of References

List of References

Adams, J., & Hitch, G. (1997). Working memory and children's mental addition. *Journal of Experimental Child Psychology*, *67*, 21-38. doi:10.1006/jecp.1997.2397.

Allport, A., Styles, E. A., & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks. In C. Umilt`a & M. Moscovitch (Eds.), *Attention and performance XV* (pp. 421–451). Cambridge, MA: MIT Press.

Alvarez, J. A., & Emory, E. (2006). Executive function and the frontal lobes: A meta-analytic review. *Neuropsychology Review, 16*, 17−42. doi:10.1007/s11065-006-9002-x.

Anderson, P. (2002). Assessment and development of executive function (EF) during childhood. *Child Neuropsychology, 8,* 71-82. doi:10.1076/chin.8.2.71.8724.

Anderson, V., Anderson, P., Northam, E., Jacobs, R., & Catroppa, C. (2001). Development of executive functions through late childhood and adolescence in an Australian sample. *Developmental Neuropsychology, 20,* 385-406. doi:10.1207/S15326942DN2001_5.

Aron, A. (2007). The neural basis of inhibition in cognitive control. *Neuroscientist, 13,* 214-228. doi:10.1177/1073858407299288.

Aron, A. (2008). Progress in executive-function research: From tasks to functions to regions to networks. *Current Directions in Psychological Science, 17,* 124-129.

Ashcraft, M. H. (1982). The development of mental arithmetic: A chronometric approach. *Developmental Review, 2,* 213-236. doi: 10.1016/0273-2297(82)90012-0

Ashcraft, M., Donley, R., Halas, M., & Vakali, M. (1992). Working memory, automaticity, and problem difficulty. In J. I. D. Campbell (Ed.), *The nature and origins of mathematical skills* (pp. 301-329). Oxford England: North-Holland.

Ashcraft, M., & Kirk, E. (2001). The relationships among working memory, math anxiety, and performance. *Journal of Experimental Psychology: General, 130,* 224-237. doi:10.1037/0096-3445.130.2.224.

Baddeley, A. D. (1996). Exploring the central executive. *Quarterly Journal of Experimental Psychology, 49,* 5–28. doi:10.1080/027249896392784.

Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends* *in Cognitive Sciences, 4,* 417-423. doi:10.1016/S1364-6613(00)01538-2.

Baddeley, A. (2002). Fractionating the central executive. In D. T. Stuss & R. T. Knight (Eds.), *Principles of frontal lobe function* (pp. 246-260). New York, NY US: Oxford University Press. doi:10.1093/acprof:oso/9780195134971.003.0016.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8, pp. 47-90). San Diego, CA: Academic Press. doi:10.1016/S0079-7421(08)60452-1.

Baldi, S., Jin, Y., Skemer, M., Green, P. J., & Herget, D. (2007). *Highlights from the PISA 2006: Performance of U.S. 15-year-old students in science and mathematics literacy in an international context* (NCES Report 2008016). Retrieved from National Center for Education Statistics website: http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2008016

Baroody, A. J., Ginsburg, H., & Waxman, B. (1983). Children's use of mathematical structure. *Journal for Research in Mathematics Education, 14,* 156–168. doi:10.2307/748379.

Baroody, A. J., & Lai, M. (2007). Preschoolers' understanding of the addition–subtraction inverse principle: A Taiwanese sample. *Mathematical Thinking & Learning, 9,* 131–171. doi:10.1080/10986060701341407.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*. 1173-1182. doi:10.1037//0022-3514.51.6.1173.

Bisanz, J., Sherman, J., Rasmussen, C., & Ho, E. (2005). Development of arithmetic skills and knowledge in preschool children. In J. I. D. Campbell (Ed.), *Handbook of Mathematical Cognition* (pp. 143–162). New York: Psychology Press.

Blair, C. (2006). How similar are fluid cognition and general intelligence? A developmental neuroscience perspective on fluid cognition as an aspect of human cognitive ability. *Behavioral and Brain Sciences, 29,* 109-160. doi:10.1017/S0140525X06009034.

Bodrova, E. & Leong, D. J. (1996). *Tools of the mind: The Vygotskian approach to early childhood education.* Upper Saddle River, NJ: Prentice-Hall.

Bogomolny, A. (1996). *Math Glossary, Math Terms, A from Interactive Mathematics Miscellany and Puzzles.* Retrieved September 19, 2010, from http://www.cut-the-knot.org/glossary/atop.shtml

Brannon, E. M. (2002). The development of ordinal numerical knowledge in infancy. *Cognition, 83*, 223–240. doi:10.1016/S0010-0277(02)00005-7.

Brocki, K., & Bohlin, G. (2004). Executive functions in children aged 6 to 13: A dimensional and developmental study. *Developmental Neuropsychology, 26*, 571-593. doi:10.1207/s15326942dn2602_3.

Bull, R., Espy, K., & Wiebe, S. (2008). Short-term memory, working memory, and executive functioning in preschoolers: Longitudinal predictors of mathematical achievement at age 7 years. *Developmental Neuropsychology, 33*, 205-228. doi:10.1080/87565640801982312.

Bull, R., & Johnston, R. (1997). Children's arithmetical difficulties: Contributions from processing speed, item identification, and short-term memory. *Journal of Experimental Child Psychology*, *65*, 1-24. doi:10.1006/jecp.1996.2358.

Bull, R., Johnston, R. S., & Roy, J. A. (1999). Exploring the roles of the visual-spatial sketch pad and central executive in children‘s arithmetical skills: Views from cognition and developmental neuropsychology. *Developmental Neuropsychology, 15*, 421–442. doi:10.1080/87565649909540759.

Bull, R., & Scerif, G. (2001). Executive functioning as a predictor of children’s mathematics ability: Inhibition, switching and working memory. *Developmental Neuropsychology, 19*, 273 293. doi:10.1207/S15326942DN1903_3.

Bunge, S. A., Dudukovic, N. M., Thomason, M. E., Vaidya, C. J., & Gabrieli, J. D. E. (2002). Immature frontal lobe contributions to cognitive control in children: Evidence from fMRI. *Neuron, 33*, 301-311. doi:10.1016/S0896-6273(01)00583-9.

Bunge, S., & Zelazo, P. (2006). A brain-based account of the development of rule use in childhood. *Current Directions in Psychological Science, 15*, 118-121. doi:10.1111/j.0963-7214.2006.00419.x.

Butterworth, B. (2005). The development of arithmetical abilities. *Journal of Child Psychology and Psychiatry, 46*, 13–18. doi:10.1111/j.1469-7610.2004.00374.x.

Butterworth, B. (1999). The mathematical brain. London: Macmillan.

Butterworth, B., Marchesini, N., & Girelli, L. (2003). Organisation of multiplication facts in memory: Developmental trends. In A. Baroody & A. Dowker (Eds.), *The development of arithmetical concepts and skills.* Mahwah, NJ: LEA.

Cantlon, J. F., Libertus, M. E., Pinel, P., Dehaene, S., Brannon, E. M., & Pelphrey, K. A. (2009). The neural development of an abstract concept of number. *Journal of Cognitive Neuroscience, 21*, 2217-2229. doi:10.1162/jocn.2008.21159.

Carlson, S., & Moses, L. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development*, *72*, 1032-1053. doi:10.1111/1467-8624.00333.

Carpenter, T. P., & Moser, J.M. (1982). The development of addition and subtraction problem solving skills. In T.P. Carpenter, J.M. Moser & T.A. Romberg (Eds.), *Addition and subtraction: A cognitive perspective* (vol. LEA, pp. 9–24). Hillsdale, NJ: LEA.

Cepeda, N., Kramer, A., & Gonzalez de Sather, J. (2001). Changes in executive control across the life span: Examination of task-switching performance. *Developmental Psychology, 37,* 715-730. doi:10.1037//0012-1649.37.5.715.

Children's Defense Fund, State of America's Children (2008). *Report of the State of America's Children 2008©.* Retrieved from http://childrensdefense.org/child-research-data-publications/data/state-of-americas-children-2008-report.pdf

Cohen, J., Cohen, P., West, S.G., & Aiken, L.S. (2003). Applied multiple regression/correlation analysis for the behavioral sciences (3rd ed.). Mahwah, NJ: Erlbaum.

Collette, F., Van der Linden, M., Laureys, S., Delfiore, G., Degueldre, C., Luxen, A., et al. (2005). Exploring the unity and diversity of the neural substrates of executive functioning. *Human Brain Mapping*, *25,* 409-423. doi:10.1002/hbm.20118.

Cowan, N. (2005). Working-memory capacity limits in a theoretical context. In C. Izawa & N. Ohta (Eds.), *Human learning and memory: Advances in theory and application: The 4th Tsukuba International Conference on Memory* (pp. 155–175). Mahwah, NJ: Erlbaum.

Crone, E., Bunge, S., van der Molen, M., & Ridderinkhof, K. (2006a). Switching between tasks and responses: A developmental study. *Developmental Science, 9*, 278-287. doi:10.1111/j.1467-7687.2006.00490.x.

Crone, E., Donohue, S., Honomichl, R., Wendelken, C., & Bunge, S. (2006b). Brain regions mediating flexible rule use during development. *Journal of Neuroscience, 26,* 11239-11247. doi:10.1523/JNEUROSCI.2165-06.2006.

Daneman, M., & Carpenter, P. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning & Verbal Behavior, 19*, 450-466. doi:10.1016/S0022-5371(80)90312-6.

De Rammelaere, S., Stuyven, E., & Vandierendonck, A. (2001). Verifying simple arithmetic sums and products: Are the phonological loop and the central executive involved? *Memory & Cognition*, *29*, 267-273.

De Smedt, B., Janssen, R., Bouwens, K., Verschaffel, L., Boets, B., & Ghesquière, P. (2009). Working memory and individual differences in mathematics achievement: A longitudinal study from first grade to second grade. *Journal of Experimental Child*

Psychology, 103, 186-201. doi:10.1016/j.jecp.2009.01.004.

Dehaene, S., Piazza, M., Pinel, P., & Cohen, L. (2003). Three parietal circuits for number processing. *Cognitive Neuropsychology, 20*, 487–506. doi:10.1080/02643290244000239.

Delazer, M., Karner, E., Zamarian, L., Donnemiller, E., Benke, T. (2006). Number processing in posterior cortical atrophy: A neuropsychological case study. *Neuropsychologia, 44*, 36–51. doi:10.1016/j.neuropsychologia.2005.04.013.

Dempster, F. (1992). The rise and fall of the inhibitory mechanism: Toward a unified theory of cognitive development and aging. *Developmental Review, 12*, 45-75. doi:10.1016/0273-2297(92)90003-K.

DeStefano, D., & LeFevre, J. (2004). The role of working memory in mental arithmetic. *European Journal of Cognitive Psychology, 16*, 353-386. doi:10.1080/09541440244000328.

Diamond, A. (1985). Development of the ability to use recall to guide action, as indicated by infants' performance on AB. *Child Development, 56*, 868-883. doi:10.1111/j.1467-8624.1985.tb00160.x.

Diamond, A. (1988). Differences between adult and infant cognition: Is the crucial variable presence or absence of language? In L. Weiskrantz (Ed.), *Thought without language* (pp. 337-370). New York: Oxford University Press.

Diamond, A. (1990). The development and neural bases of memory functions as indexed by the AB and delayed response tasks in human infants and infant monkeys. *Annals of the New York Academy of Sciences, 608*, 267-317. doi:10.1111/j.1749-6632.1990.tb48900.x.

Diamond, A. (2002). Normal development of prefrontal cortex from birth to young adulthood: Cognitive functions, anatomy, and biochemistry. In D. T. Stuss & R. T. Knight (Eds.), *Principles of frontal lobe function* (pp. 466-503). New York, NY US: Oxford University Press. doi:10.1093/acprof:oso/9780195134971.003.0029.

Diamond, A., Barnett, W. S., Thomas, J. & Munro, S. (2007). Preschool program improves cognitive control. *Science, 318*, 1387-1388. doi:10.1126/science.1151148.

Diamond, A. & Goldman-Rakic, P. S. (1985). *Evidence that maturation of frontal cortex of the brain underlies behavioral changes during the first year of life: I. The AB task.* Paper presented at the biennial meeting of the Society for Research in Child Development, Toronto.

Diamond, A., & Taylor, C. (1996). Development of an aspect of executive control: Development of the abilities to remember what I said and to 'Do as I say, not as I do.'.

Developmental Psychobiology, 29, 315-334. doi:10.1002/(SICI)1098-2302(199605)29:4<315::AID-DEV2>3.3.CO;2-C.

Dibbets, P., & Jolles, J. (2006). The switch task for children: Measuring mental flexibility in young children. *Cognitive Development, 21*, 60-71. doi:10.1016/j.cogdev.2005.09.004.

Dixon, J., Deets, J., & Bangert, A. (2001). The representations of the arithmetic operations include functional relationships. *Memory & Cognition, 29*, 462–477.

Donders, F. C. (1868/1969). On the speed of mental processes. *Acta Psychologica, 30,* 412–431. (Original work published 1868). doi:10.1016/0001-6918(69)90065-1.

Duncan, J., Emslie, H., Williams, P., Johnson, R., & Freer, C. (1996). Intelligence and the frontal lobe: The organization of goal-directed behavior. *Cognitive Psychology, 30,* 257-303. doi:10.1006/cogp.1996.0008.

Dunn, L. M., & Dunn, L. M. (2007). *Peabody Picture Vocabulary Test (4th ed.).* Bloomington, MN: Pearson Assessments.

Dunn, L. M., Dunn, L. M., Whetton, C., & Burley, J. (1997). *British Picture Vocabulary Scale, 2nd Edition.* London, UK: NFER-Nelson Publishing.

Durston, S., Thomas, K., Yang, Y., Uluğ, A., Zimmerman, R., & Casey, B. (2002). A neural basis for the development of inhibitory control. *Developmental Science, 5,* 9-16. doi:10.1111/1467-7687.00235.

Dustman, R., Emmerson, R., & Shearer, D. (1996). Life span changes in electrophysiological measures of inhibition. *Brain and Cognition, 30*, 109-126. doi:10.1006/brcg.1996.0007.

Ellefson, M. R., Blagrove, E., & Chater, N. (2010). *Developmental changes in asymmetrical switch costs*. Manuscript submitted for publication.

Ellefson, M. R., Blagrove, E., Espy, K. A., & Chater, N. (2008). *Shape school's asymmetrical switch costs*. Poster presented at the 49th annual meeting of the Psychonomic Society, Chicago, IL, USA.

Ellefson, M., Shapiro, L., & Chater, N. (2006). Asymmetrical switch costs in children. *Cognitive Development, 21*, 108-130. doi:10.1016/j.cogdev.2006.01.002.

Elliot, C. D., Smith, P., & McCullough, K. (1996). *British Ability Scales, 2nd Edition.* UK: GL Assessment Limited.

Elliot, C. D., Murray, D. J., & Pearson, L.S. (1983). *British Ability Scales*. Windsor, UK: NFER-Nelson Publishing.

Elliot, C. D. (1990). *Differential Ability Scales*. New York: The Psychological Corporation.

Espy, K. (1997). The Shape School: Assessing executive function in preschool children. *Developmental Neuropsychology*, *13*, 495-499. doi:10.1080/87565649709540690.

Espy, K., Bull, R., Martin, J., & Stroup, W. (2006). Measuring the development of executive control with the shape school. *Psychological Assessment, 18,* 373-381. doi:10.1037/1040-3590.18.4.373.

Espy, K., & Kaufmann, P. (2002). Individual differences in the development of executive function in children: Lessons from the delayed response and A-not-B tasks. In D. L. Molfese & V. J. Molfese (Eds.), *Developmental variations in learning: Applications to social, executive function, language, and reading skills* (pp. 113-137). Mahwah, NJ US: Lawrence Erlbaum Associates Publishers.

Espy, K., McDiarmid, M., Cwik, M., Stalets, M., Hamby, A., & Senn, T. (2004). The contribution of executive functions to emergent mathematic skills in preschool children. *Developmental Neuropsychology, 26,* 465-486. doi:10.1207/s15326942dn2601_6.

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics, 16,* 143-149.

Eriksen, C., & Schultz, D. (1979). Information processing in visual search: A continuous flow conception and experimental results. *Perception & Psychophysics*, *25,* 249-263.

Feigenson, L., Carey, S., & Spelke, E. (2002). Infants' discrimination of number vs. continuous extent. *Cognitive Psychology, 44,* 33-66. doi:10.1006/cogp.2001.0760.

Field, A. (2009). *Discovering statistics using SPSS: And sex and drugs and rock 'n' roll (3rd Ed).* London: Sage Publishers.

Friedman, N., Miyake, A., Corley, R., Young, S., DeFries, J., & Hewitt, J. (2006). Not all executive functions are related to intelligence. *Psychological Science*, *17,* 172-179. doi:10.1111/j.1467-9280.2006.01681.x.

Frye, D., Zelazo, P. D., & Palfai, T. (1995). Theory of mind and rule-based reasoning. *Cognitive Development, 10,* 483-527. doi:10.1016/0885-2014(95)90024-1.

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. doi:10.1037/0022-0663.98.1.29.

Fürst, A., & Hitch, G. (2000). Separate roles for executive and phonological components of working memory in mental arithmetic. *Memory & Cognition, 28*, 774-782.

Fuson, K. C. (1988). *Children's counting and concepts of number*. New York: Springer Verlag.

Gathercole, S., & Pickering, S. (2000). Working memory deficits in children with low achievements in the national curriculum at 7 years of age. *British Journal of Educational Psychology, 70,* 177-194. doi:10.1348/000709900158047.

Gathercole, S., Pickering, S., Ambridge, B., & Wearing, H. (2004a). The structure of working memory from 4 to 15 years of age. *Developmental Psychology, 40*, 177-190. doi:10.1037/0012-1649.40.2.177.

Gathercole, S., Pickering, S., Knight, C., & Stegmann, Z. (2004b). Working memory skills and educational attainment: Evidence from national curriculum assessments at 7 and 14 years of age. *Applied Cognitive Psychology, 18*, 1-16. doi:10.1002/acp.934.

Geary, D. C. (1993). Mathematical disabilities: Cognitive, neuropsychological, and genetic components. *Psychological Bulletin, 114,* 345-362. doi:10.1037/0033-2909.114.2.345.

Geary, D. (2004). Mathematics and learning disabilities. *Journal of Learning Disabilities, 37,* 4-15. doi:10.1177/00222194040370010201.

Geary, D. C., Brown, S. C., & Samaranayake, V. A. (1991). Cognitive addition: A short longitudinal study of strategy choice and speed-of-processing differences in normal and mathematically disabled children. *Developmental Psychology, 27*, 787-797. doi:10.1037/0012-1649.27.5.787.

Geary, D. C., Hamson, C. O., & Hoard, M. K. (2000). Numerical and arithmetical cognition: A longitudinal study of process and concept deficits in children with learning disability. *Journal of Experimental Child Psychology, 77*, 236-263. doi:10.1006/jecp.2000.2561.

Geier, C., Garver, K., Terwilliger, R., & Luna, B. (2009). Development of working memory maintenance. *Journal of Neurophysiology, 101*, 84-99. doi:10.1016/j.jns.2008.08.021.

Gerstadt, C., Hong, Y., & Diamond, A. (1994). The relationship between cognition and action: Performance of children 3 1/2-7 years old on a Stroop-like day-night test. *Cognition, 53*, 129-153. doi:10.1016/0010-0277(94)90068-X.

Ginsburg, H. P. (1977). *Children's Arithmetic: How They Learn It and How You Teach It.* New York, NY: Van Norwood.

Golden, C. J. (1981). The Luria-Nebraska Children's Battery: Theory and formulation. In G. W. Hynd & J. E. Obrzut (Eds.), *Neuropsychological assessment and the school-*

age child (pp. 277-302). New York: Grune & Stratton.

Gonzales, P., Williams, T., Jocelyn, L., Roey, S., Kastberg, D., Brenwald, S. (2008). *Highlights from the TIMSS 2007*(NCES Report 2009001). Retrieved from National Center for Education Statistics website: http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2005001

Gopnik, A., & Rosati, A. (2001). Duck or rabbit? Reversing ambiguous figures and understanding ambiguous representations. *Developmental Science, 4*, 175-183. doi:10.1111/1467-7687.00163.

Gordon, A., & Olson, D. (1998). The relation between acquisition of a theory of mind and the capacity to hold in mind. *Journal of Experimental Child Psychology*, *68,* 70-83. doi:10.1006/jecp.1997.2423.

Hallett, P. (1978). Primary and secondary saccades to goals defined by instructions. *Vision Research*, *18*, 1279-1296. doi:10.1016/0042-6989(78)90218-3.

Hanich, L. B., Jordan, N. C., Kaplan, D., & Dick, J. (2001). Performance across different areas of mathematical cognition in children with learning difficulties. *Journal of Educational Psychology*, *93,* 615-626. doi:10.1037/0022-0663.93.3.615.

Hitch, G. (1974). Short-term memory for spatial and temporal information. *The Quarterly Journal of Experimental Psychology*, *26*, 503-513. doi:10.1080/14640747408400440.

Hughes, M. (1986). *Children and Number: Difficulties in Learning Mathematics.* Oxford: Blackwell.

Hughes, C., & Ensor, R. (2005). Executive function and theory of mind in 2 year olds: A family affair?. *Developmental Neuropsychology*, *28*, 645-668. doi:10.1207/s15326942dn2802_5.

Huizinga, M., Dolan, C., & van der Molen, M. (2006). Age-related change in executive function: Developmental trends and a latent variable analysis. *Neuropsychologia, 44*, 2017-2036. doi:10.1016/j.neuropsychologia.2006.01.010.

Huizinga, M., & van der Molen, M. (2007). Age-group differences in set-switching and set-maintenance on the Wisconsin Card Sorting Task. *Developmental Neuropsychology*, *31*, 193-215. doi:10.1080/87565640701190817.

Huttenlocher, P. R. & Dabholkar, A. S. (1997). Developmental anatomy of prefrontal cortex. In N. Krasnegor, G. Lyon, & P. Goldman-Rakic (Eds), *Development of the prefrontal cortex: Evolution, neurobiology, and behavior* (pp. 69-83). Baltimore, MD US: Paul H Brookes Publishing.

Jacques, S., & Zelazo, P. (2005). On the possible roots of cognitive flexibility. In B. D.

Homer & C. S. Tamis-LeMonda (Eds.), *The development of social cognition and communication* (pp. 53-81). Mahwah, NJ US: Lawrence Erlbaum Associates Publishers.

Jacques, S., Zelazo, P., Kirkham, N., & Semcesen, T. (1999). Rule selection versus rule execution in preschoolers: An error-detection approach. *Developmental Psychology, 35*, 770-780. doi:10.1037//0012-1649.35.3.770.

James, W. (1890/1950). Principles of psychology. Dover Publications, New York.

Jersild, A. T. (1927). Mental set and shift. *Archives of Psychology,* Whole No. 89.

Jordan, N. C., & Hanich, L. B. (2000). Mathematical thinking in second-grade children with different forms of LD. *Journal of Learning Disabilities*, *33,* 567-578. doi:10.1177/002221940003300605.

Jordan, N., Huttenlocher, J., & Levine, S. C. (1992). Differential calculation abilities in young children from middle- and low-income families. *Developmental Psychology, 28,* 644–653. doi:10.1037//0012-1649.28.4.644.

Jordan, N. C., Levine, S. C., & Huttenlocher, J. (1995). Calculation abilities in young children with different patterns of cognitive functioning. *Journal of Learning Disabilities, 28,* 53–64. doi:10.1177/002221949502800109.

Joreskog, K. G., & Sorbom, D. (1989). LISREL7 users' reference guide. Chicago, IL: Scientific Software International.

Jöreskog, K. G. & Sörbom, D. (2009). LISREL 8.80 for Windows, Student Version [Computer Software]. Lincolnwood, IL: Scientific Software International, Inc.

Kail, R. (1991). Developmental change in speed of processing during childhood and adolescence. *Psychological Bulletin, 109,* 490-501. doi:10.1037/0033-2909.109.3.490.

Kaufmann, L. & Dowker, A. (2009). Typical development of numerical cognition: Behavioral and neurofunctional issues. *Journal of Experimental Child Psychology, 103,* 395-399. doi:10.1016/j.jecp.2009.05.003.

Kaufmann, L., Koppelstaetter, F., Siedentopf, C., Haala, I., Haberlandt, E., Zimmerhackl, et al. (2006). Neural correlates of the number-size interference task in children. *NeuroReport: For Rapid Communication of Neuroscience Research, 17*, 587–591. doi:10.1097/00001756-200604240-00007.

Karmiloff-Smith, A. (2009). Nativism versus neuroconstructivism: Rethinking the study of developmental disorders. *Developmental Psychology*, *45,* 56-63.

doi:10.1037/a0014506.

Kaye, D. B., DeWinstanley, P., Chen, Q., & Bonnefil, V. (1989). Development of efficient arithmetic computation. *Journal of Educational Psychology, 81*, 467-480. doi:10.1037/0022-0663.81.4.467.

Kelloway, E. K. (1998). Using LISREL for structural equation modeling: A researcher's guide. Thousand Oaks, CA: Sage Publications, Inc.

Kirk, U., & Kelly, M. A. (1986). Children’s differential performance on selected dorsolateral prefrontal and posterior cortical functions: A developmental perspective. *Journal of Clinical and Experimental Neuropsychology, 7*, 604.

Kirkham, N. Z., Cruess, L.,&Diamond, A. (2003). Helping children apply their knowledge to their behavior on a dimension-switching task. *Developmental Science, 6*, 449–467. doi:10.1111/1467-7687.00300.

Kline, R.B. (2005). Principles and practice of structural equation modeling. New York, NY: Guilford.

Kochanska, G., Murray, K., & Harlan, E. (2000). Effortful control in early childhood: Continuity and change, antecedents, and implications for social development. *Developmental Psychology, 36*, 220-232. doi:10.1037//0012-1649.36.2.220.

Korkman, M., Kirk, U., & Kemp, S. (1998). NEPSY: A developmental neuropsychological assessment, manual. San Antonio, TX: The Psychological Corporation.

Kray, J., Eber, J., & Lindenberger, U. (2004). Age differences in executive functioning across the lifespan: The role of verbalization in task preparation. *Acta Psychologica, 115*, 143-165. doi:10.1016/j.actpsy.2003.12.001.

Kray, J., Li, K., & Lindenberger, U. (2002). Age-related changes in task-switching components: The role of task uncertainty. *Brain and Cognition, 49*, 363-381. doi:10.1006/brcg.2001.1505.

Kroesbergen, E., Van Luit, J., Van Lieshout, E., Van Loosbroek, E., & Van de Rijt, B. (2009). Individual differences in early numeracy: The role of executive functions and subitizing. *Journal of Psychoeducational Assessment, 27*, 226-236. doi:10.1177/0734282908330586.

Kucian, K., von Aster, M., Loenneker, T., Dietrich, T., Martin, E. (2008). Development of neural networks for exact and approximate calculation: A FMRI study. *Developmental Neuropsychology, 33*, 447–473. doi:10.1080/87565640802101474.

Kuenzi, J. J. (2008). Science, technology, engineering, and mathematics (STEM) education: background, federal policy, and legislative action. *CRS Report for Congress.*

Retrieved from http://fas.org/sgp/crs/misc/RL33434.pdf

Larson, G. E., Merritt, C. R., & Williams, S. E. (1988). Information processing and intelligence: Some implications of task complexity. *Intelligence, 12*, 131–147. doi:10.1016/0160-2896(88)90012-8.

Leach, D., Coyle, C. A., & Cole, P. G. (2003). Fluency in the classroom. In R. F. Waugh (Ed), *On the forefront of educational psychology* (pp. 115-137). Hauppauge, NY: Nova Science Publishers.

Lee, K. M. (2000). Cortical areas differentially involved in multiplication and subtraction: a functional magnetic resonance imaging study and correlation with a case of selective acalculia. *Annals of Neurology, 48*, 657–661. doi:10.1002/1531-8249(200010)48:4<657::AID-ANA13>3.3.CO;2-B.

Lee, K., Ng, E., & Ng, S. (2009). The contributions of working memory and executive functioning to problem representation and solution generation in algebraic word problems. *Journal of Educational Psychology*, *101*, 373-387. doi:10.1037/a0013843.

Lehto, J., Juujärvi, P., Kooistra, L., & Pulkkinen, L. (2003). Dimensions of executive functioning: Evidence from children. *British Journal of Developmental Psychology, 21*, 59-80. doi:10.1348/026151003321164627.

Levin, H., Culhane, K., Hartmann, J., & Evankovich, K. (1991). Developmental changes in performance on tests of purported frontal lobe functioning. *Developmental Neuropsychology, 7*, 377-395. doi:10.1080/87565649109540499.

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *Journal of the Acoustical Society of America, 49*, 467-477. doi:10.1121/1.1912375.

Logan, G. (1994). On the ability to inhibit thought and action: A users' guide to the stop signal paradigm. In D. Dagenbach & T. H. Carr (Eds.), *Inhibitory processes in attention, memory, and language* (pp. 189-239). San Diego, CA US: Academic Press.

Logan, G., & Cowan, W. (1984). On the ability to inhibit thought and action: A theory of an act of control. *Psychological Review*, *91*, 295-327. doi:10.1037//0033-295X.91.3.295.

Logie, R., Gilhooly, K., & Wynn, V. (1994). Counting on working memory in arithmetic problem solving. *Memory & Cognition*, *22*, 395-410.

Lovett, M. W. (1987). A developmental approach to reading disability: Accuracy and speed criteria of normal and deficient reading skill. *Child Development*, *58*, 234-260. doi:10.2307/1130305.

Lucchelli, F. & De Renzi, E. (1993). Primary dyscalculia after a medial frontal lesion of the left hemisphere. *Journal of Neurology, Neurosurgery & Psychiatry, 56*, 304–307.

doi:10.1136/jnnp.56.3.304.

Luciana, M., Conklin, H., Hooper, C., & Yarger, R. (2005). The development of nonverbal working memory and executive control processes in adolescents. *Child Development, 76,* 697-712. doi:10.1111/j.1467-8624.2005.00872.x.

Luciana, M. & Nelson, C. (1998). The functional emergence of prefrontally-guided working memory systems in four- to eight-year-old children. *Neuropsychologia, 36,* 273-293. doi:10.1016/S0028-3932(97)00109-7.

Luna, B., Padmanabhan, A., & O'Hearn, K. (2010). What has fMRI told us about the development of cognitive control through adolescence?. *Brain and Cognition, 72,* 101-113. doi:10.1016/j.bandc.2009.08.005.

Luria, A. R. (1966). *Higher cortical function in man.* Oxford, England: Basic Books.

Luria, A. R., Pribram, K. H., & Homskaya, E. D. (1964). An experimental analysis of the behavioral disturbance produced by a left frontal arachnoidal endothelioma (meningioma). *Neuropsychologia, 2,* 257–280. doi:10.1016/0028-3932(64)90034-X.

Mayr, U. (2002). Inhibition of action rules. *Psychonomic Bulletin & Review, 9*, 93-99.

Mazzocco, M., & Kover, S. (2007). A longitudinal assessment of executive function skills and their association with math performance. *Child Neuropsychology, 13*, 18-45. doi:10.1080/09297040600611346.

McLean, J., & Hitch, G. (1999). Working memory impairments in children with specific arithmetic learning difficulties. *Journal of Experimental Child Psychology, 74*, 240-260. doi:10.1006/jecp.1999.2516.

McGuigan, N., & Núñez, M. (2006). Executive functioning by 18-24-month-old children: Effects of inhibition, working memory demands and narrative in a novel detour-reaching task. *Infant and Child Development, 15*, 519-542. doi:10.1002/icd.477.

Meiran, N. (1996). Reconfiguration of processing mode prior to task performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 1423-1442. doi:10.1037/0278-7393.22.6.1423.

Meiran, N., Gotler, A., & Perlman, A. (2001). Old age is associated with a pattern of relatively intact and relatively impaired task-set switching abilities. *The Journals of Gerontology: Series B: Psychological Sciences and Social Sciences, 56B,* 88-102.

Miller, G., Galanter, E., & Pribram, K. (1960). *Plans and the structure of behavior*. New York, NY US: Henry Holt and Co.

Milner, B. (1971). Interhemispheric differences in the localization of psychological processes in man. *British Medical Bulletin, 27*, 272-277.

Miyake, A., Friedman, N., Emerson, M., Witzki, A., & Howerter, A. (2000). The unity and diversity of executive functions and their contributions to complex 'frontal lobe' tasks: A latent variable analysis. *Cognitive Psychology, 41*, 49-100. doi:10.1006/cogp.1999.0734.

Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences, 7*, 134-140. doi:10.1016/S1364-6613(03)00028-7.

Morris, N., & Jones, D. (1990). Memory updating in working memory: The role of the central executive. *British Journal of Psychology, 81*, 111-121.

Munakata, Y. (1998). Infant perseveration and implications for object permanence theories: A PDP model of the AB task. *Developmental Science, 1*, 161-184. doi:10.1111/1467-7687.00021.

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology, 9***,** 353–383. doi:10.1016/0010-0285(77)90012-3.

Nigg, J. (2000). On inhibition/disinhibition in developmental psychopathology: Views from cognitive and personality psychology and a working inhibition taxonomy. *Psychological Bulletin, 126*, 220-246. doi:10.1037//0033-2909.126.2.220.

Oberauer, K., Süß, H., Schulze, R., Wilhelm, O., & Wittmann, W. (2000). Working memory capacity—facets of a cognitive ability construct. *Personality and Individual Differences, 29*, 1017-1045. doi:10.1016/S0191-8869(99)00251-2.

O'Leary, U., Rusch, K. M., & Guastello, S. J. (1991). Estimating age-stratified WAIS-R IQ's from scores on the Raven's Standard Progressive Matrices. *Journal of Clinical Psychology, 47*, 277-284. doi:10.1002/1097-4679(199103)47:2<277::AID-JCLP2270470215>3.0.CO;2-I.

Overman, W. H., & Bachevalier, J. (1999). Inferences about the functional development of neural systems in children via the application of animal tests of cognition. In C. A. Nelson & M. Luciana (Eds.), *Handbook of Developmental Cognitive Neuroscience* (pp. 337-357). Cambridge, MA: MIT Press

Paglin, M., & Rufolo, A. M. (1990). Heterogeneous human capital, occupational choice, and male–female earnings differences. *Journal of Labor Economics, 8,* 123–144. doi:10.1086/298239.

Passler, M. A., Isaac, W., & Hynd, G. W. (1985). Neuropsychological development of behavior attributed to frontal lobe functioning in children. *Developmental Neuropsychology, 1*, 349-370. doi:10.1080/87565648509540320.

Passolunghi, M., & Siegel, L. (2001). Short-term memory, working memory, and inhibitory control in children with difficulties in arithmetic problem solving. *Journal of Experimental Child Psychology*, *80*, 44-57. doi:10.1006/jecp.2000.2626.

Pennington, B. (1994). The working memory function of the prefrontal cortices: Implications for developmental and individual differences in cognition. In M. Haith, J. Benson, R. Roberts, B. & Pennington (Eds.), *The development of future-oriented processes* (pp. 243-289). Chicago, IL US: University of Chicago Press.

Pennington, B. (1997). Dimensions of executive functions in normal and abnormal development. In N. Krasnegor, G. Lyon, & P. Goldman-Rakic (Eds), *Development of the prefrontal cortex: Evolution, neurobiology, and behavior* (pp. 265-281). Baltimore, MD US: Paul H Brookes Publishing.

Prather, R. W., & Alibali, M. W. (2009). Development of arithmetic principle knowledge: How do we know what learners know?. *Developmental Review, 29*, 221-248. doi:10.1016/j.dr.2009.09.001.

PsychInfo®. (2010). *Database search results for "The unity and diversity of executive functions and their contributions to complex 'frontal lobe' tasks: A latent variable analysis" using the bibliolographic citation feature*. Retrieved January 1, 2010 from: http://web.ebscohost.com.proxy.library.vcu.edu/ebhost/

Ramos-Christian, V., Schleser, R., Varn, M. E. (2008). Math fluency: Accuracy versus speed in preoperational and concrete operational first and second grade children. *Early Childhood Education Journal, 35*, 543-549. doi:10.1007/s10643-008-0234-7.

Raven, J., Raven, J. C., & Court, J. H. (1993, updated 2007). *Raven's Progressive Matrices and Vocabulary Scales*. Oxford UK: Oxford Psychologists.

Ray Li, C., Huang, C., Constable, R., & Sinha, R. (2006). Imaging response inhibition in a stop-signal task: Neural correlates independent of signal monitoring and post-response processing. *The Journal of Neuroscience*, *26*, 186-192. doi:10.1523/JNEUROSCI.3741-05.2006.

Reed, M., Pien, D. L., & Rothbart, M. K. (1984). Inhibitory self-control in preschool children. *Merrill Palmer Quarterly, 30*, 131–147.

Reimers, S., & Maylor, E. (2005). Task switching across the life span: effects of age on general and specific switch costs. *Developmental Psychology, 41*, 661-671. doi:10.1037/0012-1649.41.4.661.

Ridderinkhof, K., & van der Molen, M. (1995). A psychophysiological analysis of developmental differences in the ability to resist interference. *Child Development*, *66*, 1040-1056. doi:10.2307/1131797.

Rivera, S. M., Reiss, A. L., Eckert, M. A., & Menon, V. (2005). Developmental changes in mental arithmetic: Evidence for increased functional specialization in the left inferior parietal cortex. *Cerebral Cortex*, *24*, 50–60. doi:10.1093/cercor/bhi055.

Rivera-Batiz, F. L. (1992). Quantitative literacy and the likelihood of employment among young adults in the United States. *Journal of Human Resources, 27*, 313–328. doi:10.2307/145737.

Rogers, R. D., & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General, 124***,** 207–231. doi:10.1037//0096-3445.124.2.207.

Salthouse, T., Babcock, R., & Shaw, R. (1991). Effects of adult age on structural and operational capacities in working memory. *Psychology and Aging*, *6*, 118-127. doi:10.1037//0882-7974.6.1.118.

Schneider,W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime user's guide*. Pittsburgh, PA: Psychology Software Tools, Inc.

Scope, A., Empson, J., & McHale, S. (2010). Executive function in children with high and low attentional skills: Correspondences between behavioural and cognitive profiles. *The British Journal of Developmental Psychology*, *28*, 293-305. doi:10.1348/026151009X410371.

Semenza, C., Miceli, L., Girelli, L. (1997). A deficit for arithmetical procedures: Lack of knowledge or lack of monitoring?. *Cortex*, *33*, 483–498. doi:10.1016/S0010-9452(08)70231-4.

Senn, T., Espy, K., & Kaufmann, P. (2004). Using path analysis to understand executive function organization in preschool children. *Developmental Neuropsychology*, *26*, 445-464. doi:10.1207/s15326942dn2601_5.

Seyler, D., Kirk, E., & Ashcraft, M. (2003). Elementary subtraction. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 1339-1352. doi:10.1037/0278-7393.29.6.1339.

Shallice, T. (1982). Specific impairments in planning. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences*, *298*, 199–209. doi:10.1098/rstb.1982.0082.

Sherman, J., & Bisanz, J. (2007). Evidence for use of mathematical inversion by three-year-old children. *Journal of Cognition & Development*, *8*, 333–344. doi:10.1080/15248370701446798.

Siegler, R. S., & Booth, J. (2005). Development of numerical estimation: a review. In J. I. D. Campbell (Ed.), *Handbook of Mathematical Cognition* (pp. 197–212). New York:

Psychology Press.

Smith-Chant, B. L., & LeFevre, J. (2003). Doing as they are told and telling it like it is: Self-reports in mental arithmetic. *Memory and Cognition, 31*, 516-528.

Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. In S. Leinhardt (Ed.), *Sociological Methodology 1982* (pp. 290-312). Washington DC: American Sociological Association.

Sobel, M. E. (1988). Direct and indirect effects in linear structural equation models. In J. S. Long (Ed.), *Common Problems/Proper Solutions: Avoiding Error in Quantitative Research* (pp. 46-64). Beverly Hills, CA: Sage Publishers.

Spector, A., & Biederman, I. (1976). Mental set and mental shift revisited. *American Journal of Psychology, 89***,** 669–679. doi:10.2307/1421465.

Squire, S., Davies, C., & Bryant, P. (2004). Does the cue help? Children's understanding of multiplicative concepts in different problem contexts. *British Journal of Educational Psychology, 74,* 515–532. doi:10.1348/0007099042376364.

St Clair-Thompson, H., & Gathercole, S. (2006). Executive functions and achievements in school: Shifting, updating, inhibition, and working memory. *The Quarterly Journal of Experimental Psychology, 59*, 745-759. doi:10.1080/17470210500162854.

Stanescu-Cosson, R., Pinel, P., van De Moortele, P. F., Le Bihan, D., Cohen, L., & Dehaene, S. (2000). Understanding dissociations in dyscalculia a brain imaging study of the impact of number size on the cerebral networks for exact and approximate calculation. *Brain; a Journal of Neurology, 123*, 2240-2255. doi:10.1093/brain/123.11.2240.

Starkey, P., & Cooper, R.G. (1980). Perception of numbers by human infants. *Science, 210*, 1033–1035. doi:10.1126/science.7434014.

Starkey, P., & Gelman, R. (1982). The development of addition and subtraction abilities prior to formal schooling in arithmetic. In T.P. Carpenter, J.M. Moser & T.A. Romberg (Eds.), *Addition and subtraction: A cognitive perspective* (pp. 99–116). Hillsdale, NJ: LEA.

Starkey, P., Spelke, E.S., & Gelman, R. (1990). Numerical abstraction by human infants. *Cognition, 36*, 97–128. doi:10.1016/0010-0277(90)90001-Z.

Strommen, E. A. (1973). Verbal self-regulation in a children's game: Impulsive errors on "Simon says." *Child Development, 44*, 849–853. doi:10.1111/j.1467-8624.1973.tb01165.x.

Stroop, J. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental*

Psychology, 18, 643-662. doi:10.1037/h0054651.

Swanson, H. L., & Beebe-Frankenberger, M. (2004). The relationship between working memory and mathematical problem solving in children at risk and not at risk for serious math difficulties. *Journal of Educational Psychology, 96*, 471-491. doi:10.1037/0022-0663.96.3.471.

Swanson, H., Jerman, O., & Zheng, X. (2008). Growth in working memory and mathematical problem solving in children at risk and not at risk for serious math difficulties. *Journal of Educational Psychology, 100*, 343-379. doi:10.1037/0022-0663.100.2.343.

Swanson, H., & Sachse-Lee, C. (2001). Mathematical problem solving and working memory in children with learning disabilities: Both executive and phonological processes are important. *Journal of Experimental Child Psychology, 79*, 294-321. doi:10.1006/jecp.2000.2587.

Tamm, L., Menon, V., & Reiss, A. (2002). Maturation of brain function associated with response inhibition. *Journal of the American Academy of Child & Adolescent Psychiatry*, *41*, 1231-1238. doi:10.1097/00004583-200210000-00013.

Thomas, M., & Karmiloff-Smith, A. (2002). Are developmental disorders like cases of adult brain damage? Implications from connectionist modelling? *Behavioral and Brain Sciences*, *25*, 727-787. doi:10.1017/S0140525X02000134.

Thorndike, R. L., Hagen, E. P., & Sattler, J. M. (1986). *Stanford-binet intelligence scales*. Chicago, IL: Riverside Publishing Co.

Tolar, T. D., Lederberg, A. R. & Fletcher, J. M. (2009). A structural model of algebra achievement: Computational fluency and spatial visualisation as mediators of the effect of working memory on algebra achievement. *Educational Psychology, 29*, 239-266. doi:10.1080/01443410802708903.

Trbovich, P., & LeFevre, J. (2003). Phonological and visual working memory in mental addition. *Memory & Cognition*, *31*, 738-745.

van Boxtel, G., van der Molen, M., Jennings, J., & Brunia, C. (2001). A psychophysiological analysis of inhibitory motor control in the stop-signal paradigm. *Biological Psychology, 58*, 229-262. doi:10.1016/S0301-0511(01)00117-X.

Van Dooren, W., De Bock, D., Hessels, A., Janssens, D., & Verschaffel, L. (2005). Not everything is proportional: Effects of problem type and age on propensities for over-generalization. *Cognition and Instruction, 23*, 57-86. doi:10.1207/s1532690xci2301_3.

Van Dooren, W., De Bock, D., & Verschaffel, L. (2010). From addition to multiplication…

and back: The development of students' additive and multiplicative reasoning skills. *Cognition and Instruction, 28,* 360-381. doi:10.1080/07370008.2010.488306.

van der Sluis, S., de Jong, P., & van der Leij, A. (2007). Executive functioning in children, and its relations with reasoning, reading, and arithmetic. *Intelligence, 35*, 427-449. doi:10.1016/j.intell.2006.09.001.

Van Harskamp, N. J., Cipolotti, L. (2001). Selective impairments for addition, subtraction and multiplication: Implications for the organisation of arithmetical facts. *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior*, *37*, 363–388. doi: 10.1016/S0010-9452(08)70579-3.

Van Loosbroek, E., & Smitsman, A.W. (1990). Visual perception of numerosity in infancy. *Developmental Psychology, 26*, 916–922. doi:10.1037/0012-1649.26.6.911.b.

Vock, M., & Holling, H. (2008). The measurement of visuo-spatial and verbal-numerical working memory: Development of IRT-based scales. *Intelligence*, *36*, 161-182. doi:10.1016/j.intell.2007.02.004.

Wallis, J., & Miller, E. (2003). Neuronal activity in primate dorsolateral and orbital prefrontal cortex during performance of a reward preference task. *European Journal of Neuroscience, 18,* 2069-2081. doi:10.1046/j.1460-9568.2003.02922.x.

Warrington, E. K. (1982). The fractionation of arithmetical skills: A single case study. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 34,* 31–51. doi:10.1080/14640748208400856.

Waugh, N., & Norman, D. (1965). Primary memory. *Psychological Review, 72,* 89-104. doi:10.1037/h0021797.

Wechsler, D. (1949). *The Wechsler Intelligence Scale for Children*. New York: Psychological Corp.

Wechsler, D. (1981). *Wechsler Adult Intelligence Scale-Revised.* New York: The Psychological Corporation.

Wechsler, D. (1991). *Wechsler Intelligence Scale for Children – Third Edition*. San Antonio: Psychological Corp.

Wechsler, D. (2003). *Wechsler Intelligence Scale for Children—Fourth Edition.* San Antonio, TX: Harcourt.

Wechsler, D. (2008). *Wechsler Adult Intelligence Scale—Fourth Edition.* San Antonio, TX: Pearson Assessment.

Welsh, M., Friedman, S., & Spieker, S. (2006). Executive functions in developing children:

Current conceptualizations and questions for the future. In K. McCartney & D. Phillips (Eds), *Blackwell Handbook of Early Childhood Development* (pp. 167-187). Malden: Blackwell Publishing. doi:10.1002/9780470757703.ch9.

Welsh, M., & Pennington, B. (1988). Assessing frontal lobe functioning in children: Views from developmental psychology. *Developmental Neuropsychology, 4*, 199-230. doi:10.1080/87565648809540405.

Welsh, M., Pennington, B., & Groisser, D. (1991). A normative-developmental study of executive function: A window on prefrontal function in children. *Developmental Neuropsychology, 7*, 131-149. doi:10.1080/87565649109540483.

Wilkins, J., Baroody, A. J., & Tiilikainen, S. (2001). Kindergartners' understanding of additive commutativity within the context of word problems. *Journal of Experimental Child Psychology, 79*, 23–36. oi:10.1006/jecp.2000.2580.

Wilkinson, G. S. (1993). *Wide Range Achievement Test, 3rd ed*. Lutz, FL: Psychological Assessment Resources, Inc.

Wynn, K. (1992). Addition and subtraction by human infants. *Nature, 358*, 749-750. doi:10.1038/358749a0.

Xu, F., & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition, 74*, B1-B11. doi:10.1016/S0010-0277(99)00066-9.

Yntema, D. B. (1963). Keeping track of several things at once. *Human Factors, 5*, 7−17.

Zamarian, L., Ischebeck, A., Delazer, M. (2009). Neuroscience of learning arithmetic: Evidence from brain imaging studies. *Neuroscience and Biobehavioral Reviews, 33*, 909–925. doi:10.1016/j.neubiorev.2009.03.005.

Zelazo, P., Craik, F., & Booth, L. (2004). Executive function across the life span. *Acta Psychologica, 115*, 167-183. doi:10.1016/j.actpsy.2003.12.005.

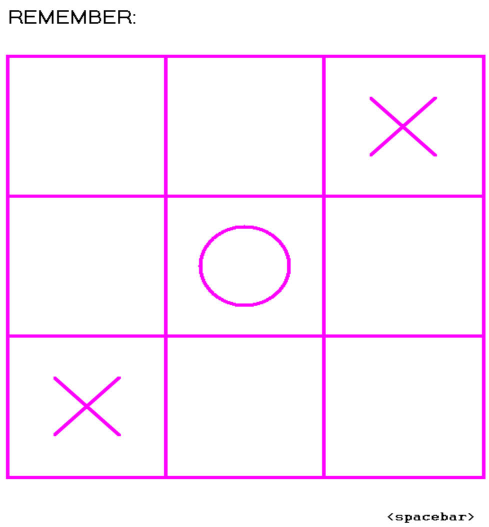
Zelazo, P., Müller, U., Frye, D., & Marcovitch, S. (2003). The development of executive function. *Monographs of the Society for Research in Child Development, 68*, 127. doi:10.1111/j.0037-976X.2003.00261.x.

Zelazo, P. D., Reznick, J. S., & Piñon, D. E. (1995). Response control and the execution of verbal rules. *Developmental Psychology 31*, 508-517. doi:10.1037//0012-1649.31.3.508.

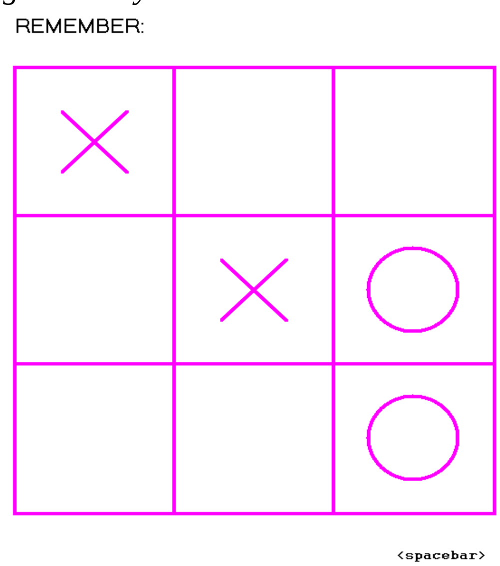
Appendix A

Tic-Tac-Toe
Sample stimulus displays from the Tic-Tac-Toe task (Huizinga et al., 2006; Milner, 1971) included as part in the test battery (used with permission, Ellefson et al., 2010)

Low Memory Load Trial



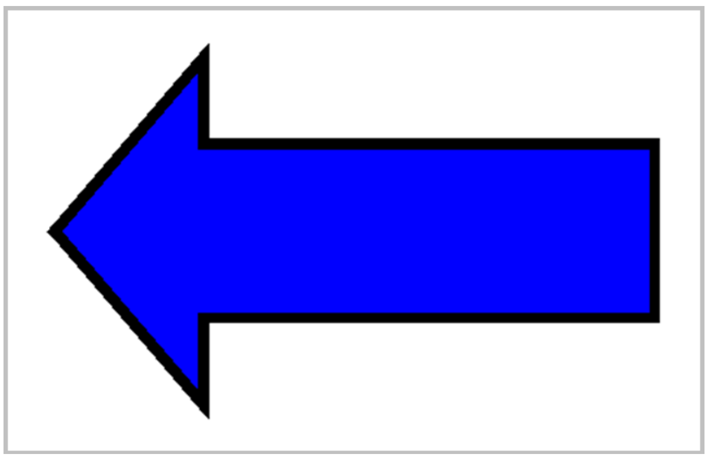
High Memory Load Trial



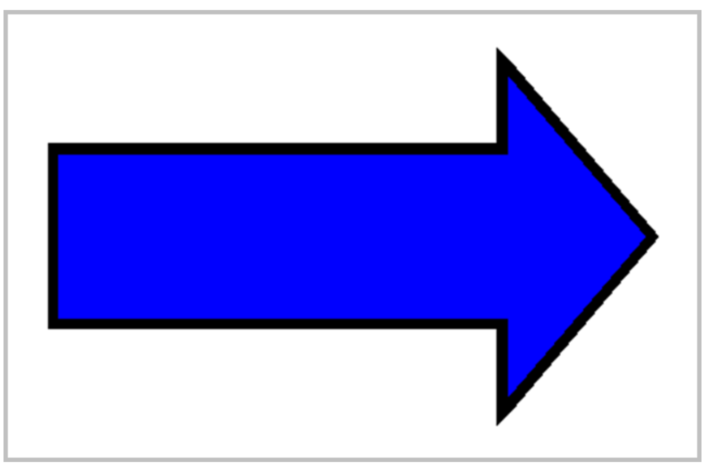
Appendix B

Stop Signal
Sample stimulus displays from the Stop Signal task (Logan, 1994) included as part in the test battery (used with permission, Ellefson et al., 2010).

Left Trial

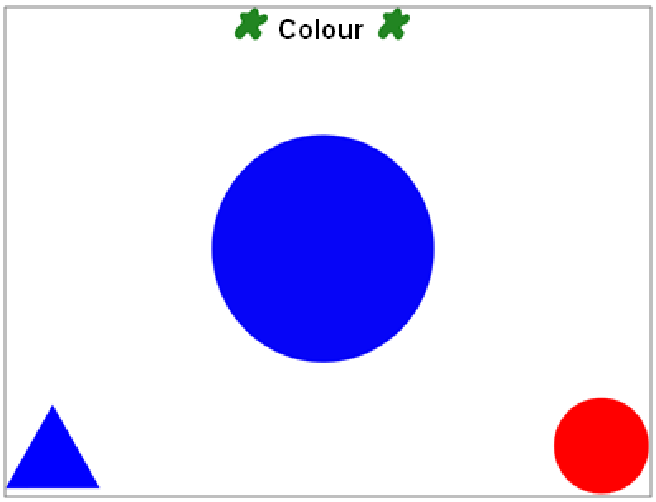


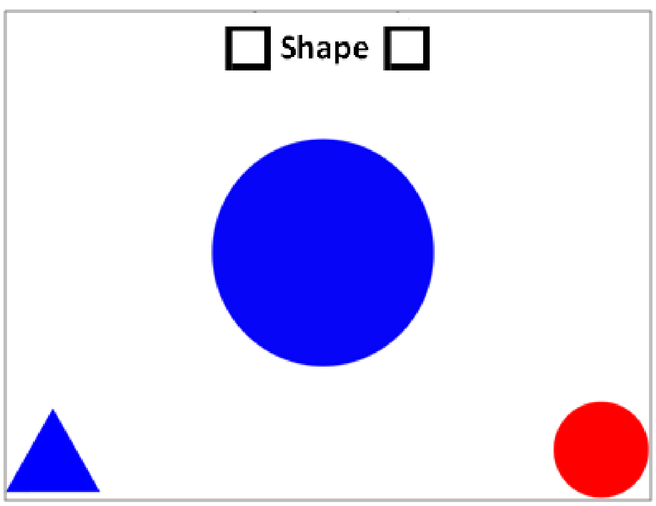
Right Trial



Appendix C

Figure Matching
Sample stimulus displays from the Figure Matching task (Ellefson et al., 2006) included as part in the test battery (used with permission, Ellefson et al., 2010).

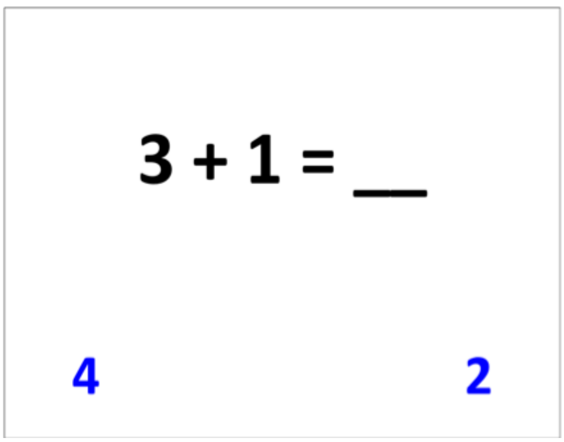


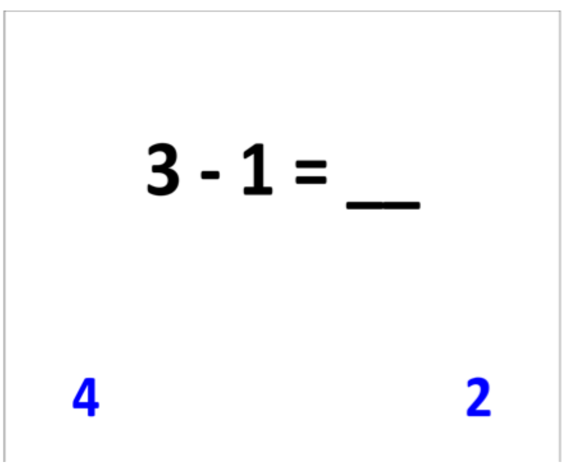


Appendix D

Arithmetic Matching

Sample stimulus displays from the Arithmetic Matching task (Ellefson et al., 2006) included as part in the test battery (used with permission, Ellefson et al., 2010).





Appendix E

Missing data calculations

Missing Variable	Instances	Predictor Variable	Regression Equation
TTT Effic	2	Digits (Raw)	7.211^{e-5} (Digits) + .000
Digits (Raw)	1	TTT Effic	3246.69 (TTT Effic) + 18.55
Figures Log Effic	2	SS-Sw Effic	66.146 (SS-Sw Effic) + 249.802
SS-SW Effic	1	Figures Log Effic	0.005 (Figures Log Effic) + -0.911
Log SSRT	1	SS-Inh Log Effic	0.369 (SS-Inh Log Effic) + 2.347

Note. Effic = Efficiency; Log = Logarithmic transformed.

Appendix F

Hierarchical Regression Models Summary Table: Sets 3-7

Variable	R^2	Adj R^2	ΔR^2	ΔF(df)	*p*
MATH Composite Model 3					
Step 1: AGE	.576	.573	.576	198.00 (1, 146)	< .001
Step 2: GCA-RAVENS, GCA-BPVS	.695	.688	.119	28.08 (2, 144)	< .001
Step 3: WM-DIGITS, WM-TTT	.719	.709	.024	6.13 (2, 142)	.003
Step 4: SS-SS-SW, SS-FIGURES	.749	.736	.030	8.27 (2, 140)	< .001
Step 5: INH-SS-INH, INH-SSRT	.768	.753	.019	5.80 (2, 138)	.004
MATH-WRAT Model 3					
Step 1: AGE	.566	.563	.566	190.67 (1, 146)	< .001
Step 2: GCA-RAVENS, GCA-BPVS	.779	.774	.213	69.33 (2, 144)	< .001
Step 3: WM-DIGITS, WM-TTT	.803	.796	.023	8.43 (2, 142)	< .001
Step 4: SS-SS-SW, SS-FIGURES	.825	.816	.023	9.05 (2, 140)	< .001
Step 5: INH-SS-INH, INH-SSRT	.833	.823	.008	3.46 (2, 138)	.034
MATH Composite Model 4					
Step 1: AGE	.576	.573	.576	198.00 (1, 146)	< .001
Step 2: GCA-RAVENS, GCA-BPVS	.695	.688	.119	28.08 (2, 144)	< .001
Step 3: INH-SS-INH, INH-SSRT	.740	.731	.046	12.45 (2, 142)	< .001
Step 4: WM-DIGITS, WM-TTT	.753	.740	.013	3.55 (2, 140)	.031
Step 5: SS-SS-SW, SS-FIGURES	.768	.753	.015	4.58 (2, 138)	.012
MATH-WRAT Model 4					
Step 1: AGE	.566	.563	.566	190.67 (1, 146)	< .001
Step 2: GCA-RAVENS, GCA-BPVS	.779	.774	.213	69.33 (2, 144)	< .001
Step 3: INH-SS-INH, INH-SSRT	.808	.801	.029	10.57 (2, 142)	< .001
Step 4: WM-DIGITS, WM-TTT	.820	.811	.012	4.61 (2, 140)	.012

Variable	R^2	Adj R^2	ΔR^2	ΔF(df)	*p*
Step 5: SS-SS-SW, SS-FIGURES	.833	.823	.014	5.76 (2, 138)	.004
MATH-TSMATH Model 4					
Step 1: AGE	.416	.412	.416	104.19 (1, 146)	< .001
Step 2: GCA-RAVENS, GCA-BPVS	.451	.439	.034	4.51 (2, 144)	.013
Step 3: INH-SS-INH, INH-SSRT	.502	.484	.051	7.27 (2, 142)	.001
Step 4: WM-DIGITS, WM-TTT	.517	.493	.015	2.20 (2, 140)	.114
Step 5: SS-SS-SW, SS-FIGURES	.531	.500	.014	2.01 (2, 138)	.139
MATH Composite Model 5					
Step 1: AGE	.576	.573	.576	198.00 (1, 146)	< .001
Step 2: GCA-RAVENS, GCA-BPVS	.695	.688	.119	28.08 (2, 144)	< .001
Step 3: INH-SS-INH, INH-SSRT	.740	.731	.046	12.45 (2, 142)	< .001
Step 4: SS-SS-SW, SS-FIGURES	.757	.745	.017	4.88 (2, 140)	.009
Step 5: WM-DIGITS, WM-TTT	.768	.753	.011	3.27 (2, 138)	.041
MATH-WRAT Model 5					
Step 1: AGE	.566	.563	.566	190.67 (1, 146)	< .001
Step 2: GCA-RAVENS, GCA-BPVS	.779	.774	.213	69.33 (2, 144)	< .001
Step 3: INH-SS-INH, INH-SSRT	.808	.801	.029	10.57 (2, 142)	< .001
Step 4: SS-SS-SW, SS-FIGURES	.825	.817	.018	7.11 (2, 140)	.001
Step 5: WM-DIGITS, WM-TTT	.833	.823	.008	3.33 (2, 138)	.039
MATH-TSMATH Model 5					
Step 1: AGE	.416	.412	.416	104.19 (1, 146)	< .001
Step 2: GCA-RAVENS, GCA-BPVS	.451	.439	.034	4.51 (2, 144)	.013
Step 3: INH-SS-INH, INH-SSRT	.502	.484	.051	7.27 (2, 142)	.001
Step 4: SS-SS-SW, SS-FIGURES	.514	.490	.012	1.77 (2, 140)	.175
Step 5: WM-DIGITS, WM-TTT	.531	.500	.017	2.44 (2, 138)	.091
MATH Composite Model 6					
Step 1: AGE	.576	.573	.576	198.00 (1, 146)	< .001

Variable	R^2	Adj R^2	ΔR^2	ΔF(df)	*p*
Step 2: GCA-RAVENS, GCA-BPVS	.695	.688	.119	28.08 (2, 144)	< .001
Step 3: SS-SS-SW, SS-FIGURES	.734	.725	.040	10.57 (2, 142)	< .001
Step 4: WM-DIGITS, WM-TTT	.749	.736	.014	4.01 (2, 140)	.020
Step 5: INH-SS-INH, INH-SSRT	.768	.753	.019	5.80 (2, 138)	.004
MATH-WRAT Model 6					
Step 1: AGE	.566	.563	.566	190.67 (1, 146)	< .001
Step 2: GCA-RAVENS, GCA-BPVS	.779	.774	.213	69.33 (2, 144)	< .001
Step 3: SS-SS-SW, SS-FIGURES	.813	.807	.034	12.97 (2, 142)	< .001
Step 4: WM-DIGITS, WM-TTT	.825	.816	.012	4.77 (2, 140)	.010
Step 5: INH-SS-INH, INH-SSRT	.833	.823	.008	3.46 (2, 138)	.034
MATH-TSMATH Model 6					
Step 1: AGE	.416	.412	.416	104.19 (1, 146)	< .001
Step 2: GCA-RAVENS, GCA-BPVS	.451	.439	.034	4.51 (2, 144)	.013
Step 3: SS-SS-SW, SS-FIGURES	.484	.466	.034	4.63 (2, 142)	.011
Step 4: WM-DIGITS, WM-TTT	.501	.476	.017	2.36 (2, 140)	.098
Step 5: INH-SS-INH, INH-SSRT	.531	.500	.029	4.33 (2, 138)	.015
MATH Composite Model 7					
Step 1: AGE	.576	.573	.576	198.00 (1, 146)	< .001
Step 2: GCA-RAVENS, GCA-BPVS	.695	.688	.119	28.08 (2, 144)	< .001
Step 3: SS-SS-SW, SS-FIGURES	.734	.725	.040	10.57 (2, 142)	< .001
Step 4: INH-SS-INH, INH-SSRT	.757	.745	.023	6.60 (2, 140)	.002
Step 5: WM-DIGITS, WM-TTT	.768	.753	.011	3.27 (2, 138)	.041
MATH-WRAT Model 7					
Step 1: AGE	.566	.563	.566	190.67 (1, 146)	< .001
Step 2: GCA-RAVENS, GCA-BPVS	.779	.774	.213	69.33 (2, 144)	< .001
Step 3: SS-SS-SW, SS-FIGURES	.813	.807	.034	12.97 (2, 142)	< .001
Step 4: INH-SS-INH, INH-SSRT	.825	.817	.012	4.91 (2, 140)	.009
Step 5: WM-DIGITS, WM-TTT	.833	.823	.008	3.33 (2, 138)	.039

Variable	R^2	Adj R^2	ΔR^2	ΔF(df)	*p*
MATH-TSMATH Model 7					
Step 1: AGE	.416	.412	.416	104.19 (1, 146)	< .001
Step 2: GCA-RAVENS, GCA-BPVS	.451	.439	.034	4.51 (2, 144)	.013
Step 3: SS-SS-SW, SS-FIGURES	.484	.466	.034	4.63 (2, 142)	.011
Step 4: INH-SS-INH, INH-SSRT	.514	.490	.030	4.28 (2, 140)	.016
Step 5: WM-DIGITS, WM-TTT	.531	.500	.017	2.44 (2, 138)	.091

Vita

Andrea Renee Molzhon was born on September 6, 1984, in Midlothian, Virginia, and is an American citizen. She graduated from Midlothian High School, Midlothian, Virginia in 2002. She received her Bachelor of Science in Psychology from Longwood University, Farmville, Virginia in 2006 and subsequently attended one semester of graduate school in the Community Counseling program at George Washington University, Washington, D.C. She began the doctoral program in Applied Developmental Psychology at Virginia Commonwealth University in August, 2008. She has been working as a Graduate Research Assistant on an Early Reading First grant with the School of Education at Virginia Commonwealth University, Richmond, VA, from August 2009 to present.