MATH FLUENCY DEVELOPMENT IN THE EARLY ELEMENTARY GRADES AND THE

ROLE OF THE SUMMER BREAK, INATTENTION, AND WORKING MEMORY

By

Danielle V.E. Pigon

A thesis submitted in conformity with the requirements
for the degree of Doctor of Philosophy
Graduate Department of Applied Psychology and Human Development Ontario Institute for
Studies in Education of the
University of Toronto

**MATH FLUENCY DEVELOPMENT IN THE EARLY ELEMENTARY GRADES AND**

**THE ROLE OF THE SUMMER BREAK, INATTENTION, AND WORKING MEMORY**

Doctor of Philosophy, 2017
Danielle V.E. Pigon
Department of Applied Psychology and Human Development University of Toronto

**Abstract**

The current dissertation examined factors thought to influence the developmental dynamics of basic math fluency skill (addition and subtraction). The overarching goal of this research was to examine growth patterns in addition and subtraction fluency across Grades 1 through 4 while taking into account the effects of the non-instructional period over the summer months. This was accomplished using an accelerated longitudinal design within the framework of hierarchical linear modeling. A cumulative pattern of growth was observed over time for each fluency measure. Differences between addition and subtraction growth trajectories were found and are discussed.

This dissertation also aimed to examine the respective roles of verbal working memory, visual-spatial working memory, and teacher ratings of classroom inattention in the development of math fluency while controlling for sex and parent level of education. This was done from a piecewise perspective, that is, by considering whether these cognitive and behavioral factors exerted differential effects based on seasonal patterns (i.e., skill growth that occurs within the school year (excluding the summer), and the growth that occurs across grades (including the summer)). Differential seasonal patterns were identified: Visual-spatial working memory was predictive of skill development specific to the school year, whereas teacher-rated inattention was a key longitudinal predictor of weaker math fluency, across the years, including the summer months. Implications for research and educational practice are discussed.

## Acknowledgements

I would like to express my sincere gratitude to my supervisor, Dr. Rosemary Tannock. Thank you, Dr. Tannock, for being exceptionally generous with your time and knowledge. I am especially appreciative of your patience, encouragement, guidance, and kindness, which helped me regain the confidence I needed to keep moving forward when the road was feeling particularly long and arduous. You are a mentor in the truest sense, and it has been a pleasure and a privilege to be your trainee.

Thank you to my committee members, Dr. Rhonda Martinussen and Dr. Richard Wolfe, for being so giving of your time and expertise, and for your enthusiasm for this project. Special thanks to Dr. Martinussen for your structural suggestions that helped pull the final document together, and to Dr. Wolfe for opening up a world of statistics that I never could have dreamed of, and making it not only meaningful, but exciting.

To my Internal Examiner, Dr. Michele Peterson-Badali: Thank you for being an integral part of my graduate school journey, from my very first assessment course, right up to the Final Oral Examination. It has been an absolute pleasure being your student over the years.

Thank you to my External Examiner, Dr. Jo-Anne LeFevre. Your contributions during the final stages of this work were instrumental in helping me continue to think both broadly and deeply about my research.

Finally, to my family: Mom, Dad, Paul, and Patrick. Thank you for always finding the perfect balance between accepting where I am, and encouraging me to keep moving toward my goals. This one goal, especially, would not have been possible without your unconditional love and support; I will forever be grateful.

# Table of Contents

# List of Tables

# List of Figures

# List of Appendices

# CHAPTER I
## Overview and Rationale of the Dissertation Research

**Overview of the Dissertation**

Research on children's mathematical skill has seen a significant surge over the past two decades. A number of these studies have utilized longitudinal designs. Critical to developmental research, longitudinal designs provide an advantage over cross-sectional studies as they help identify the nature of change over time, allow for the assessment of the variability and patterns of this change, and the identification of factors associated with change. Under certain circumstances, they can even contribute qualified causal inferences about cause and effect relationships (Little, Card, Preacher, & McConnell, 2009). In addition to gaining a better understanding of human development, from an educational perspective, information gleaned from these studies can inform theories of academic achievement as well as decisions regarding both practice and policy.

Much of the extant longitudinal research on the developmental trajectories of math comes from large-scale national studies. Many of these studies range a number of years in childhood and include thousands of participants, such as the Early Childhood Longitudinal Study – Kindergarten Cohort (ECLS-K) (Judge & Watson, 2011; Kohli, Sullivan, Sadeh, & Zopluoglu, 2015; Lee, 2010; Li-Grining, Votruba-Drzal, Maldonado-Carreno, & Haas, 2010; Morgan, Farkas, & Wu, 2009; Morgan, Farkas, & Wu, 2011; Schulte & Stevens, 2015) and the National Longitudinal Survey of Youth (NLSY) (Duncan et al., 2007; Leahey & Guo, 2001; Stipek & Valentino, 2015). However, most of these studies have approached the domain of mathematics from a general perspective, simultaneously tapping a variety of mathematical abilities (e.g., conceptual knowledge, procedural knowledge, problem-solving) reflecting grade-appropriate content. The importance of these studies in the advancement of our understanding of the nature of math development is undeniable, as they provide important insight into overall math development at an epidemiological level. However, one particular challenge to the study of mathematics is its heterogeneous nature; therefore, taking a general approach may obscure

developmental variability in the acquisition of the various sub-skills. These sub-skills may have differing developmental trajectories, may be associated with different sets of predictive factors, and may lead to diverging long-term effects on achievement (Fuchs et al., 2006; Geary, 1993; Jordan, Hanich, & Kaplan, 2003; Vasilyeva, Laski, & Shen, 2015). For example, and most relevant to the current study, children may experience difficulties in basic math fluency (i.e., the ability to solve basic arithmetic facts with both speed and accuracy), despite normal achievement in other areas of math, such as story problems (Hanich, Jordan, Kaplan, & Dick, 2001; Jordan et al., 2003).  Therefore, it is equally important to have insight into the developmental trajectories of specific math sub-skills. To this end, the current dissertation focuses on the development of the sub-skill of math fluency, which has seen relatively little research attention, particularly from a longitudinal perspective using community samples.

**Organization of the dissertation**. The current dissertation focuses on the development of math fluency in the early elementary grades. It is organized in 4 chapters. The first chapter is comprised of a comprehensive literature review and delineation of the purpose of the study. The second chapter outlines the methods used in the dissertation research, where growth of basic math fluency across Grades 1 through 4 is examined from a longitudinal perspective while taking into account the effect of the summer months, thereby considering both between-grade (across years) and within-grade (within school years) skill growth. This study also investigates the roles of verbal (numerical) working memory, visual-spatial working memory, and teacher-rated inattention on the development of math fluency while controlling for the effects of sex and parent level of education. Results are presented in the third chapter. The fourth chapter is a comprehensive discussion of the findings, as well as implications for research and practice. The Appendix includes further details regarding methodology (scoring outlines).

**Literature Review**

**Influence of Quantitative Literacy Across the Lifespan.**

Poor arithmetic skills are considered a risk factor for occupational and health outcomes in adulthood, as quantitative literacy is a strong predictor of high school completion, employment status, occupational earnings, and health-related decision-making in adulthood in both industrialized and developing nations, over and above the influence of literacy, educational level, and intelligence (Boissiere, Knight, & Sabot, 1985; Golbeck, Ahlers-Schmidt, Paschal, & Dismuke, 2005; Montori & Rothman, 2005; Parsons & Bryner, 2005; Rivera-Batiz, 1992). Math difficulties that appear early (e.g., at school entry, around ages 5 or 6) have been identified as an important risk factor for weaknesses in math at the end of schooling (Duncan et al., 2007; Pagani, Fitzpatrick, Archambault, & Janosz, 2010) regardless of family background, social-emotional factors, reading skill, and intelligence. Further, previous research estimates that between 5-8% of school-age children experience some form of mathematics learning difficulty (Geary, 2004). Considering the impact of mathematical competency across the lifespan and the current movement in educational practice and psycho-educational service delivery models towards prevention and early intervention (Meyers & Nastasi, 1999; Reisener, Dufrene, Clark, Olmi, & Tingstrom, 2016), gaining a clearer understanding of the developmental dynamics of mathematical ability as well as the elucidation of factors that contribute to the developmental variability is critical. The current dissertation examines the developmental dynamics of the foundational sub-skill of math fluency (Fuchs et al., 2006), and investigates the role of factors that may predict individual differences in the context of normative development.

**Math Fluency**

      **Definition and development of math fluency.** A number of studies have examined the development of basic arithmetic in terms of accuracy, or the ability to obtain a correct response regardless of time (Aunola, Leskinen, Lerkkanen, & Nurmi, 2004; Morgan et al., 2009; Morgan et al., 2011). A related, though distinct sub-skill is math fluency; in the context of the current dissertation[1], math fluency refers to the narrow ability of computational proficiency, or the ability to recall (or efficiently calculate) simple arithmetic problems ("math facts") with both speed and accuracy (Binder, 1996; Clarke, Nelson, & Shanley, 2016). For *mastery with fluency* to occur, the accurate recall (or use of other procedures, see below) of math facts must be accompanied by conceptual knowledge (Baroody, Bajwa, & Eiland, 2009), which promotes the efficient and adaptive use of facts and procedures (e.g., knowing whether or not algorithm can be used) within the context of both familiar and novel tasks (Baroody et al., 2009; Clarke et al., 2016). The development of math fluency has received notably less research attention than accuracy, despite being recognized as an important focus for mathematics instruction as highlighted by educational standards in the U.S. (National Council of Teachers of Mathematics, 2000; National Mathematics Advisory Panel, 2008). Although the importance of fluency is noted in the current Ontario[2] mathematics curriculum (Ontario Ministry of Education, 2005), explicit acquisition through memorization is deemphasized. The Ontario curriculum adopts a discovery-based method of teaching, where a greater importance is placed on a conceptual understanding of mathematics than on memorization, or on students' ability to follow a "correct" procedure to a known answer. This is somewhat in contrast to Québec, which also follows a discovery-based

---

[1] As summarized by Clarke et al. (2016), math fluency can also be conceptualized as a broad construct that relates to having gained proficiency in various math domains (e.g., early numeracy, whole number concepts, rational number concepts, and algebra).

[2] Students in the current study were attending school in Ontario, Canada.

curriculum, albeit with an integrated requirement of memorization of basic math facts and algorithms (Gouvernement du Québec, Ministère de l'Éducation, 2001).

Most research related to math fluency development has been conducted from a cognitive lens, and primarily relating to the development of increasingly sophisticated strategies used to solve basic arithmetic problems (Bailey, Littlefield, & Geary, 2012; Geary, Hoard, & Nugent, 2012; Vasilyeva et al., 2015). From this perspective, math fluency develops as children gradually become less reliant on quantity-based procedural strategies, such as counting, and increasingly proficient at flexibly using a mix of fact retrieval from long-term memory and back-up problem-solving strategies (e.g., sophisticated and/or rapid counting and decomposition) (Bailey et al., 2012; De Smedt, Taylor, Archibald, & Ansari, 2010; Geary, BowThomas, Liu, & Siegler, 1996; Geary et al., 2012; Mazzocco, Devlin, & McKenney, 2008; Shrager & Siegler, 1998; Siegler & Shipley, 1995; Siegler & Stern, 1998; Thevenot, Barrouillet, Castel, & Uittenhove, 2016). Models of skill acquisition demonstrate how practice supports the transition through distinct learning states leading to increased automaticity (Tenison & Anderson, 2015). In the case of arithmetic, solving simple arithmetic through counting leads to repeated exposure to a problem and its answer, which in turn leads to the formation of increasingly strong associations between problem and answer pairs (Shrager & Siegler, 1998; Siegler & Shrager, 1984). The probability of direct retrieval from memory is dependent on the strength of this association, which itself increases with practice; therefore, retrieval becomes increasingly frequent with practice (Imbo, Vandierendonck, & Rosseel, 2007) and advancing development (Siegler & Shipley, 1995). This has been suggested for addition (Geary & Burlinghamdubree, 1989) and subtraction (Siegler, 1987). Fluency is also related to problem size, such that problems involving smaller numbers (e.g., 2 + 4) are solved more quickly and accurately than problems involving larger numbers (e.g., 7 + 9); this *problem size effect* has been observed in both adults (e.g., LeFevre et al., 1996) and children (De Smedt, Holloway, & Ansari, 2011; Imbo & Vandierendonck, 2008). In

children, this effect is related to the finding that smaller problems are more likely to be retrieved directly from memory, whereas larger problems are more likely to be solved through procedural strategies (Imbo & Vandierendonck, 2008). By the mid-elementary school years (10-12 years), children who tend to rely more on quantity-based strategies are less fluent in simple arithmetic, compared to peers who tend to retrieve facts directly from memory (De Smedt et al., 2011). However, because a mix of retrieval and procedural strategies continue to be used throughout the lifespan (De Smedt et al., 2010; Fayol & Thevenot, 2012; Imbo et al., 2007; LeFevre, DeStefano, Penner-Wilger, & Daley, 2006; LeFevre et al., 1996), math fluency is related both to the frequency and speed of basic fact retrieval from long-term memory and to the sophistication, accuracy, and speed of backup problem-solving procedures, such as counting and decomposition (Geary et al., 1996; Geary et al., 2012; Mazzocco et al., 2008; Siegler & Stern, 1998; Vasilyeva et al., 2015).

**Effects of math fluency ability and difficulty**. Increased fluency has been linked to important educational outcomes. Generally, fluency contributes to improved retention and maintenance of learned material, increased task endurance and attention span, greater resistance to distraction, and the ability to apply what has been learned to the performance of higher level skills (Binder, 1996; Lindsley, 1996). Fluency is a critical foundation of math development, as operations that are automatized require limited allocation of attentional or other cognitive resources (e.g., working memory) (Ackerman, Anhalt, & Dykman, 1986; Binder, 1996; Imbo & Vandierendonck, 2007; Lindsley, 1996), thus freeing up these resources, which can then be allocated to the acquisition of more complex skills needed to tackle advanced aspects of problem solving in childhood (Carr & Alexeev, 2011; Carr, Taasoobshirazi, Stroud, & Royer, 2011; Fuchs et al., 2006; Hecht, Torgesen, Wagner, & Rashotte, 2001; Vasilyeva et al., 2015) and adolescence (Price, Mazzocco, & Ansari, 2013). In this way, as argued by Gersten and Chard

(1999), a lack of arithmetic fluency may present a barrier to understanding higher-level aspects of the curriculum. Indeed, previous studies have found that simple fact fluency (i.e., single digit operations) and strategy choice are foundational to the acquisition of computation skills in terms of accuracy (Vasilyeva et al., 2015), as well as word-problem solving ability (Fuchs et al., 2006; Hecht et al., 2001). The notion that fluency scaffolds higher level math abilities has also been supported by a neuroimaging study which demonstrated that activation in functional brain networks known to be associated with fact retrieval (i.e., left supramarginal gyrus and bilateral anterior cingulate cortex) during a simple arithmetic task was related to higher scores on a Grade 10 students' Preliminary Scholastic Aptitude Test (a standardized test used to predict college entrance exams) scores (Price et al., 2013). Conversely, greater activation in the right hemisphere regions known to be associated with processing of quantity (right intraparietal sulcus) during the arithmetic task was associated with weaker PSAT scores. A further possible advantage to achieving fluent performance includes a link to lower levels of math anxiety (Cates & Rhymer, 2003).

Conversely, consistent research findings have shown that deficits in fact fluency are hallmark features of both math difficulties and disabilities, regardless of the presence of comorbid reading disabilities (Geary, Hoard, & Bailey, 2012; Jordan, Hanich, & Kaplan, 2003). In fact, it has been proposed that it is the specific failure in becoming fluent in basic math facts that underlies many students' mathematical difficulties (Bull & Johnston, 1997; Geary, Brown, & Samaranayake, 1991).

**Math fluency as dissociable from accuracy.** Measures of math fluency and accuracy may appear similar (e.g., assessing basic arithmetic), with the key difference being that speed is taken into account in the scoring and interpretation of the former. Converging lines of evidence suggest that the construct of math fluency is dissociable from that of untimed basic arithmetic.

8

From a mastery perspective, prior work has indicated that taking performance rate into account provides a more sensitive gauge of skill growth and proficiency than accuracy alone (Ackerman et al., 1986; Deno, 1985; Prindle, Mitchell, & Pretscher, 2016; Shapiro, DuPaul, Bradley, & Bailey, 1996; Zentall, 1990). Conversely, scores from untimed tests may mask differences in ability (Ackerman et al., 1986; Binder, 2003; Ramos-Christian, Schleser, & Varn, 2008). For example, in an early study investigating the children's transition from the use of counting to retrieval from memory, Ashcraft and Fierman (1982) noted that without time constraints, children in third grade were able to complete simple addition as accurately as children in Grade 6, despite greater fluency in the latter group. Similarly, Jordan and Montani (1997) found that children with math disabilities were able to correctly complete computation tasks in terms of accuracy when given enough time, despite weakness in fluency. Binder, Haughton, and Bateman (2002) noted the pitfalls of a "percentage correct world" where 100% represents the highest level of performance. Their assertion that perfect accuracy is not the definition of mastery can be highlighted using a simple example: consider two children who are both able to perform at 100% accuracy on a math task; the first child is able to give 10 correct responses in an allotted time, whereas the second child completes 5 within the same time frame. Given time-based information, it could be argued that the second child does not know the material as well as the first. However, this information would not be reflected in their accuracy scores.

Behavioral studies that have considered arithmetic both in terms of accuracy and fluency have shown that math fluency accounts for unique variance in math ability over and above that of untimed math skill (Fuchs et al., 2008; Hart, Petrill, & Thompson, 2010; Mazzocco et al., 2008). In a 2011 study, Carr and Alexeev examined children's math development; although gender, single-digit accuracy and single digit fluency each exerted an influence on higher-level math ability, fluency had the most significant impact on math skill growth. From a genetic perspective, studies with school-aged twins identified math fluency as being etiologically distinct

from untimed basic arithmetic, untimed reading, and reading fluency (Hart, Petrill, Thompson, & Plomin, 2009; Hart et al., 2010; Petrill et al., 2012). Further, with respect to measurement, in a study of students in second to fifth grade, Burns and colleagues (2006) identified that the categorization of students into instructional levels (i.e., frustration, instructional, and mastery levels) that enable the matching of students' ability (based on their baseline fluency skills) to appropriate instructional approaches, was significantly more reliable using math fluency compared to accuracy data. These same authors found that fluency outcomes were more closely related to criterion measures (standardized general math outcomes) compared to accuracy data. Considering the converging evidence for the distinction between timed and untimed skill, it cannot be assumed that fluency development mirrors the development of basic arithmetic skill in terms of accuracy.

**Longitudinal development of math fluency.** When considering studies that have investigated the developmental trajectories of math fluency, most come from the math learning disability literature. These studies compare groups of students who experience math difficulties and/or disability to their typically achieving peers (Chong & Siegel, 2008; Geary et al., 1991; Geary et al., 2012; Jordan, Kaplan, & Hanich, 2002; Jordan et al., 2003; Ostad, 1997; Vanbinst, Ghesquiere, & De Smedt, 2014). Note that in the literature, the convention is generally that children who score at or below the 10th percentile on standardized tests of math achievement over at least two consecutive years are considered to fall within the *math learning disability* category, whereas children who score between the 11th and 25th percentile are categorized as having *learning difficulties* (Geary, 2011b), although some authors include children up to the 35th percentile (Jordan et al., 2003).

Overall, these studies identify important differences in terms of the development of fluent math skill across subgroups of ability. For example, research examining fluency development

from the perspective of forced retrieval from memory tasks has shown that children with relatively weaker fluency skill (defined as falling below the 35th percentile) experienced virtually no growth over Grades 2-3 despite growth in other areas of mathematics, such as broad math and story problems (Jordan et al., 2003). This was in contrast to students with stronger fluency skills who demonstrated linear growth within the same timeframe. In another study, stable weakness in fluency (direct retrieval measured across Grades 2 and 3) was found to be a feature of children with math disability, regardless of whether this weakness occurred alone (MD-only), or with a comorbid reading disability (MD-RD) (Jordan et al., 2003). In contrast, children who were typically achieving (TA) and those with a reading disability (RD-only) had higher slopes. Growth rates were similar for all four groups, indicating that the low fluency groups did not catch up to groups with stronger fluency, over time (Jordan et al., 2003). Similarly, Chong and Siegel (2008) found that low-achieving children or those with a math learning disability experienced overall weaker fluency over Grades 2-5, compared to typically achieving students. In contrast to the findings of Jordan and colleagues (2003), however, Chong and Siegel observed a linear increase in fluency growth within the study's timeframe for all three groups, and no between-group differences were observed in terms of rate of fluency increase. Sampling over a longer timeframe (across Grades 1 to 5), Geary and colleagues (2012) found that fact retrieval fluency and decomposition skill development followed a curvilinear (decelerating) trajectory over time in typically developing and persistently low achieving children; the quadratic slope effect was non-significant in the math disabled group. Further, there were significant differences in the mean levels of the slopes in each group. Other authors have investigated variability in longitudinal profile types (Jordan, Kaplan, Olah, & Locuniak, 2006). For example, recently, VanBinst and colleagues (2014) used a data-driven, model-based clustering approach to delineate profiles of arithmetic fact fluency based on repeated fluency assessment at the beginning of third, fourth, and fifth grades. These authors found three distinct profiles which

were comparable in terms of age, sex, SES, and intellectual ability: slow and variable, average, and efficient.

These studies are of course essential to the enhancement of our understanding of the nature, development, course, and correlates of math learning disability. However, group based-approaches do not provide a sense of person-level variability across the spectrum of ability. Further to this point, Geary and colleagues (2012) did not find distinct MLD and LA groups when using cluster and growth curve analyses, indicating that children composing these groups are part of the normal distribution of mathematics achievement.

Of the studies that have evaluated fluency longitudinally from a person-centered general education perspective, between-year comparisons have generally been made from a group (grade) perspective. For example, in an earlier study, Fuchs and colleagues (1993) examined students' rates of growth in math fluency using mathematics curriculum-based measurement (M-CBM) in a community sample of students in Grades 1-6 to determine the weekly rate of student progress in fluency. Two samples were collected over two consecutive years (Year 1, $n = 177$ and Year-2, $n = 1,208$ for math measures). Students were assessed weekly in Year 1 and at least once per month in Year 2. The authors conducted person-centered analyses by calculating individual slopes for each student using OLS regression, and the distributions of slopes were examined. Based on a randomly selected subset (n = up to 56) at each grade level, linearity was determined to represent the functional form within the grades (although a sizeable minority of Grade 6 students showed a decelerating pattern of growth within the grade). Cross-year analyses were performed to examine the relationship between the slope and grade level, using one-way ANOVA. This analysis revealed gradual deceleration in the magnitude of improvement with advancing years.

Building on the study by Fuchs et al. (1993), Graney and colleagues (2009) examined the within-year development of math fluency (using Mathematics-CBM), for three grade cohorts

(Grade 3, 4, and 5) for two consecutive years. Data were collected in the fall, winter, and spring of each year, and was analyzed using repeated measures ANOVA. Similar to the Fuchs et al. (1993) study, the work of Graney and colleagues showed that growth rates increased with successive grade levels; however, they also found greater growth between winter and spring, compared to the growth that occurred from fall to winter. In a more recent study building on Graney and colleagues' work (2009), Keller-Margulis, Mercer, and Shapiro (2014) also found evidence for non-linearity in within-year math growth as assessed by M-CBM, with greater growth from fall to winter compared to winter to spring. These authors noted that increased growth in the fall may have been reflective of a rapid recoupment of skills lost over the summer, once skills are reviewed at the return of school in the fall (Allinder & Eicher, 1994; Keller-Margulis et al., 2014).

Martens, Hurks, Meijs, Wassenberg, and Jolles (2011) investigated differences between children's (age 6-15 years) fluency in the four operation types (addition, subtraction, multiplication, and division). Using general linear modeling, they noted improvement in each operation type across the years. The largest improvements occurred yearly until Grade 5, with continued (though less) improvement until grade 8. Further, these authors found that fluency level varied according to operation types, as further discussed below. More recently, using a large sample (N = 4337), Van de Weijer-Bergsma, Kroesbergen, and Van Luit (2015) also examined fluency performance according to operation types (i.e., addition, subtraction, multiplication, and division). Students from Grades 2 to 6 were tested longitudinally at 3 time points over the course of one school year. Multilevel multi-group latent growth modeling was used to examine differences in level and growth with advancing Grades. These authors also examined the predictive value of verbal and visual-spatial working memory, which will be discussed further below. Results showed that for all operation types, within-grade growth increased with advancing grades, although the magnitude of this growth decreased in the highest

13

grades assessed (Grades 5 and 6). The authors also found increased variability between student levels and rates of growth in the early grades, leading to greater variability over time. However, this gap appeared to stabilize in the later grades, as rates of growth became increasingly comparable despite variable levels of mean ability.

Collectively, relatively little is known about the development and variability of math fluency across the spectrum of ability. Further, in longitudinal studies that have taken a person-level approach to the study of math fluency in typically developing children, between-grade analyses have been performed at a group level. Therefore, studies investigating developmental growth trajectories of general stream education students both from a within- and between-grade perspective are lacking, as highlighted by both Fuchs et al. (1993) and Graney et al. (2009). Based on the research, there is little doubt that problems with math fluency are most evident in children with persistent math disabilities or difficulties. However, studies also highlight the need for research in the normative development of basic fluency. First, as mentioned previously, fluency in basic arithmetic is fundamental to higher level math skills (Price et al., 2013). This is contrasted with evidence that typically achieving children are themselves not fluent in basic arithmetic (Henry & Brown, 2008). For example, in a study on 275 children in first grade, Henry and Brown (2008) found that children (including those from high performing schools) achieved the state standard of direct retrieval of sums (and corresponding differences) to 18 only 22% of the time. This was despite there being a specific focus on this goal in California schools at the time of the study. In Canada, skill in this area appears to be declining over time (LeFevre, Penner-Wilger, Pyke, Shanahan, & Deslauriers, 2014). Recently, LeFevre and colleagues (2014) identified a staggering decline (by 23%) in university undergraduates' basic arithmetic fluency that has occurred gradually over the past 20 years. These findings were based on data from university students from 1993 and 2005, who began receiving formal arithmetic instruction between 1982 and 1993 (i.e., in their Grade 1 year). The authors noted further decline, estimated

at 37%, for data collected between 2005 and 2010. LeFevre and colleagues hypothesized that these findings might be linked to significant changes in the curriculum over the past 20 years, where fluency has been deemphasized in the context of more discovery-based instructional methods, a general expansion in math areas of instruction, as well as increased calculator use. This concerning trend highlights the need to have a clear understating of fluency development regardless of disability level, with a focus on individual variability.

## Developmental Dynamics of Math Ability

**Functional Form of Math Growth Trajectories.** With respect to modeling growth, smooth linear and curvilinear models are almost universally seen in development and education literature. Accurately modeling the functional form (shape) of the growth trajectory can provide insight into the nature and developmental stages in a given academic domain. Specifically, in the presence of consistent practice, the acquisition of skilled behavior is often presumed to follow a relatively smooth curvilinear (or sigmoid, S-shaped) form, with a period of rapid growth occurring at a consistent rate, followed by a gradual slowing of growth in skill slowing over time (Speelman & Kirsner, 2005). Gradual slowing in skill acquisition might occur as an individual approaches the optimal level of skilled performance (e.g., mastery). Alternatively, skill slowing may be seen if practice is halted, or if cognitive resources are redirected to other (e.g., higher level) tasks (e.g., Speelman & Kirsner, 2005).

Most researchers studying the growth of general math ability through childhood have noted a curvilinear pattern (J. Lee, 2010; Mok, McInerney, Zhu, & Or, 2015; Morgan et al., 2009; Murayama, Pekrun, Lichtenfeld, & Vom Hofe, 2013; Muthen & Khoo, 1998; Schulte & Stevens, 2015; T. Shin, Davison, Long, Chan, & Heistad, 2013), although others have identified linear growth (Rescorla & Rosenthal, 2004). For general mathematics, linear growth may be reflective of its cumulative nature (i.e., new concepts are built upon previously learned ones)

(Clarke et al., 2106), whereas curvilinear growth may indicate that material becomes increasingly challenging with advancing grades (Mok et al., 2015; Shin et al., 2013). As outlined in the previous section, in terms of math fluency, linear growth has been generally identified in shorter-term longitudinal studies (e.g., those capturing within-grade growth (e.g., Fuchs et al., 1993)), or in studies capturing growth across earlier elementary grades (e.g., Grades 2-3 as seen in Jordan et al., 2003a and 2003b). Curvilinear trends might be deduced in studies examining longer time-spans, which have generally identified decreased growth with increasing grade levels (Fuchs et al., 1993; Graney et al., 2009; Martens et al., 2011; Van de Weijer-Bergsma et al., 2015). However, as noted above, these between-grade changes were examined at a group rather than at the individual level; the benefits of using a growth curve modeling approach over a repeated measures (Singer & Willett, 2003) are highlighted in the next section.

Of course, discrepancies in findings regarding the form of the growth curve may be in part due to differences in study variables, including variability in the math subcomponents assessed, characteristics of the student sample, differences in measures used (see Graney et al., 2009), whether the timeframe of the study adequately captures the form of the trajectory (e.g., earlier years when a skill is still being acquired, versus later years), the frequency of data collection, and the timing of data collection (e.g., fall vs. spring). Therefore, an important consideration to understanding the form of the developmental trajectory using a longitudinal design is measuring the skill during the time of developmental change (Little et al., 2009). In Ontario, teaching standards indicate that addition and subtraction are a focus of the Grade 1 curriculum (Ontario Ministry of Education, 2005). Previous studies have suggested that children gradually transition from counting-based strategies to memory-based strategies around the third grade (e.g., Ashcraft & Fierman, 1982), although they continue to use procedural-based strategies in 3rd and 4th grade (Barrouillet, Mignon, & Thevenot, 2008; Robinson, 2001). According to the National Mathematics Advisory Panel (2008), in the U.S., achievement of

proficiency in basic addition and subtraction is a goal to be reached by the end of third grade. The current study spans the Grades 1 through 4, thus capturing theoretically critical time-points within the development of basic arithmetic fluency, and with an adequate number of time points (i.e., 8) to model various functional forms of growth.

**Discontinuous/piecewise models of growth and the summer slowdown.** An alternative to smooth growth models are piecewise ones, which propose that growth occurs in distinct phases with their own distinct shapes. Longitudinally, this creates a jagged trajectory, in contrast to the smooth trajectories seen in linear or curvilinear models. Few researchers in the area of math have adopted a piecewise modeling strategy, although studies that have utilized it have found that it provides a more adequate description of general math development than curvilinear models (Kohli et al., 2015; Shin et al., 2013). For example, in a recent study using U.S. national longitudinal data from the Early Childhood Longitudinal Study (Kindergarten Cohort; ECLS-K), Kohli and colleagues (2015) found that a piecewise trajectory most accurately represented growth in children's (Grade K-8) mathematics skills as compared to other trajectory forms. Notably, they also observed slowed growth rates between the spring of Kindergarten and the fall of Grade 1 (i.e., the two grades in which data had been collected in the fall and spring of each year). However, subsequently, data was collected only once per year in the spring of third, fifth, and eighth grades, which did not allow them to detect similar between-year variation in higher grades. In the realm of fluency, earlier studies also point to a piecewise development, although these effects were not directly examined. For example, Fuchs et al. (1993) noted that while a linear relationship adequately captured within-year growth of children's math fluency skill, growth across Grades 1 to 6 was curvilinear. However, Fuchs and colleagues acknowledged that they were unable to simultaneously model both within- and between-year growth because the within-year analyses were conducted at the student level, whereas between-year analyses were

conducted at a group level. Similarly, as noted above, Graney and colleagues (2009) studied within-year growth for Grades 3, 4, and 5, for two consecutive years; however, their analyses did not permit them to simultaneously investigate the growth across the grades (between year) from a longitudinal perspective. Notably, curvilinear functions may appear to be linear in a restricted range (e.g., within a single grade), which may in part explain why studies observe linear growth within a grade, and curvilinear trajectory when examining changes in grade means (e.g., Fuchs et al., 1993). Piecewise functions are capable of approximating linear or curvilinear functions, thus offering a suitable option to address this design issue. Conditions necessary for conducting a piecewise analysis include: (a) that data are obtained from the same students in at least two different grades, (b) that the dependent variable has a consistent scale at all time points, and (c) that the analysis considers variability in growth trajectories. A piece-wise analysis would not have been possible in the Fuchs and colleagues' (1993) study, as although they used data from two consecutive years, the sample sizes were different (year 1, $n = 177$; year 2, $n = 1,208$). In Graney and colleagues' (2009) study, a subset of students would have participated over two years (i.e., those in Grade 3 to 4, and Grade 4 to 5). However, their analysis consisted of group comparison using an ANOVA approach, which did not permit analysis of variability in growth trajectories. Keller-Margulis and colleagues (2014) were also unable to conduct such an analysis as their data was only collected in one year for students in Grades 1 to 5, although as mentioned above, they noted the effects of the summer as a potential explanation for the observed differences in with higher rates of growth in the fall and decelerating rates of growth over the spring (linked to rapid recoupment after the summer months). The current research builds on these previous studies by utilizing a piecewise analysis to simultaneously estimate growth from both a within-year and between-year perspective.

Given that children learn most of their math skills within the context of the school environment, an important variable to consider in the measurement of math development is the

timing of data collection within the school year (e.g., fall vs. spring). Generally, we would expect differences based on the timing of the assessment, such that assessments performed early in the year reflects skills learned in the previous year, whereas assessments administered later in the school year would also reflect learning that occurred within that grade level (Anderman, Gimbert, O'Connell, & Riegel, 2015). The importance of assessment timing is further highlighted by a body of educational research demonstrating that long breaks in instruction over the summer affect rates of academic skill growth (reading and math) over time (Alexander, Entwisle, & Olson, 2001; Cooper, Nye, Charlton, Lindsay, & Greathouse, 1996; Davies & Aurini, 2013; Downey, von Hippel, & Broh, 2004). This slowing of skill acquisition or skill loss during an instruction-free period is referred to as the *summer slide* (Alexander et al., 2007; Vale, 2013) or *summer slowdown (*Vale et al., 2013). Other terms used to describe this phenomenon include *summer setback* (Davies & Aurini, 2013; Entwistle & Alexander, 1992; Vale, 2013; Downey et al., 2004), *summer learning loss (*Patton & Reschly, 2013), and *summer effects (*Cooper et al., 1996). These terms are used interchangeably in the literature; the term *summer slowdown* will be used in the current thesis.  Theoretically, a piecewise model to growth in math fluency skill across the grades is relevant in light of the summer slowdown literature (Cooper et al., 1996; Davies & Aurini, 2013; Downey et al., 2004; Vale et al., 2013), as the peaks and valleys in learning rates reflecting seasonal changes (i.e., within the school year and during the summer) could not be adequately captured using a single linear or curvilinear function.

Most studies considering the summer break have focused on the effects of within-school versus out of school (e.g., home and neighborhood) variables. These studies have demonstrated a link between SES and the development of learning gaps, particularly in literacy, and mainly considering U.S. schools (Alexander et al., 2001; Alexander, Entwisle, & Olson, 2007a; Alexander, Entwisle, & Olson, 2007b), although Canadian (Davies & Aurini, 2013) and Australian (Vale et al., 2013) studies noted similar findings. Studies examining the effects of the

summer slowdown have shown that mathematical ability, and computation in particular (i.e., as opposed conceptual mathematics, such as problem-solving), is highly vulnerable to skill loss or slowing of skill acquisition over the summer months (Allinder, Fuchs, Fuchs, & Hamlett, 1992; Allinder & Eicher, 1994; Cooper et al., 1996). Some studies have found that children from disadvantaged backgrounds make fewer gains in the summer and begin the new school year lagging behind their higher-SES peers and that this pattern is cumulative over the years (Alexander et al., 2001; Alexander et al., 2007a; Downey et al., 2004; Verachtert, Van Damme, Onghena, & Ghesquiere, 2009). Other studies have not found a significant difference in the gains made by high and low SES groups, with all children experiencing significant loss in mathematics (Cooper et al., 1996). In a meta-analysis of the summer effects literature, Cooper and colleagues (1996) found that children were losing the equivalent to approximately 2.6 months of computation skill over the summer holiday. Notably, these authors found that this effect occurred regardless of children's sex, IQ, race, and family income level.  Cooper and colleagues proposed that the loss of skill is particularly profound in math calculation because procedural math skills are susceptible to forgetting in the absence of practice, and because math is not generally practiced over the summer months (Cooper et al., 1996; Cooper & Sweller, 1987). Although studies that have investigated the effects of the summer months on math achievement have typically done so by considering general math or computation in terms of accuracy (e.g., Downey et al., 2004), similar effects on fluency skill have been documented (Allinder et al., 1992; Allinder & Eicher, 1994). Considering the fact that practice is a significant factor in fluency growth (e.g., Tenison & Anderson, 2015), loss or slowing in skill growth over the summer months would be unsurprising. One might further postulate that loss of math fluency skill specifically, may be linked to the fact that speed of retrieval increases with practice (Bailey et al., 2012; Royer, Tronsky, Chan, Jackson, & Marchant, 1999).

The effect of the summer slowdown suggests that even when studies of varying time-spans identify similar functional forms (e.g., linear within a grade and linear across the grades), the estimated rates of growth, within and across grades, cannot be assumed to be equivalent. Collectively, the studies outlined above suggest that the structure of the school year is an important environmental factor to consider when examining developmental growth, and once-yearly measurement is likely insufficient to adequately capture growth patterns. In other words, although the development of math fluency may follow a general linear or curvilinear trajectory across a number of years, this trajectory likely also contains within-grade components. Adopting a piecewise modeling strategy would allow for the explicit parceling-out the time period during which children are not receiving instruction (Anderson, 2012; Downey et al., 2004; Verachtert et al., 2009). To our knowledge, this hypothesis has not been empirically examined with respect to math fluency. The piecewise modeling approach used in the current study allows for the simultaneous examination of growth patterns that are specific to the school year (i.e., excluding the summer) as well as across Grades 1 through 4 (including the summers).

**Cumulative versus compensatory growth patterns.** Although is well established that early math ability is a strong predictor of proficiency in later mathematics (Duncan et al., 2007; Pagani et al., 2010), the developmental process by which children attain this proficiency is unclear. When examining the developmental of individual differences in growth trajectories, an important consideration involves the relationship between children's early math ability and the rate at which their skills develop. In the presence of a significant correlation between initial skill level and growth rate, two primary views describe the process through which children gain proficiency in math ability (Ackerman, 2007; Aunola et al., 2004); specifically, the acquisition of math skill can occur in either a cumulative or compensatory fashion (Aunola et al., 2004; Mok et al., 2015; Morgan et al., 2009; Morgan et al., 2011; Muthen & Khoo, 1998; Shin et al., 2013). In

the case of cumulative growth, ability increases through a process of cumulative advantage or disadvantage; that is, children who start out as having stronger basic skills (e.g., acquired through informal learning of mathematics concepts during the preschool years) continue to acquire skills more rapidly, whereas children who begin with weaker basic math skills have slower growth rates. Longitudinally, this pattern leads to a "fan-spread" effect (Aarnoutse & van Leeuwe, 2000), with increasing variability between students' ability as they advance through the grades, and growing achievement gaps between those with strong and those with weak initial abilities, with weaker students fall increasingly behind (Ackerman, 2007; Walberg & Tsai, 1983). This is also referred to as the Matthew effect, widely recognized in the reading literature (Aarnoutse & van Leeuwe, 2000; Stanovich, 1986). A cumulative growth pattern when measuring general math ability may be reflective of the hierarchical nature of math skills. For example, difficulty understanding whole numbers will be a barrier to understating fractions, which can subsequently lead to difficulties tackling higher-level math such as algebra (National Mathematics Advisory Panel, 2008). Cumulative effects may be linked to losses occurring over the summer months, establishing widening gaps between certain subgroups of children (e.g., along SES lines) (Alexander et al., 2001; Alexander et al., 2007a). In the case of math fluency, a cumulative cycle may be linked to proficiency in math fluency itself (e.g., fluency leads to greater fluency). For example, considering that practice is an important predictor of fluency (Binder, 1996), children who are more fluent would have more opportunity to practice a greater number of items in an allocated time (more completion of items), which could theoretically lead to greater fluency. Some authors have demonstrated that stronger fluency affects children's maturation of strategy choice, by supporting a transition from the use of manipulatives to mental problem-solving approaches (Carr & Alexeev, 2011). Cumulative effects may also be linked to underlying cognitive abilities such as working memory (Bailey, Watts, Littlefield, & Geary, 2014; Geary et al., 2012), basic number competency (Jordan, Kaplan, Ramineni, & Locuniak, 2009), or

behavioural traits such as approaches to learning and student engagement (Bodovski & Farkas, 2007; Li-Grining et al., 2010; McClelland, Acock, & Morrison, 2006).

An alternative pattern is a compensatory one, which supposes that children who initially have weaker math skills, eventually "catch up" to their peers. This could be the case for lower skilled children who enter school at an academic or social disadvantage, but who may benefit from receiving systematic instruction upon entering the school system (which, as highlighted by Aunola and colleagues (2004), has been the case in reading (Phillips, Norris, Osmond, & Maynard, 2002)), advantageous school characteristics (e.g., school safety, qualified teachers, adequate resources (Han, 2008)), becoming increasingly engaged with instruction (Bodovski & Farkas 2007), or receiving effective interventions (Ramani & Siegler, 2008; Siegler, 2009). Finally, in the case where the correlation between the intercepts of the growth trajectory and the rates of growth is zero, gaps remain stable across the years (Jordan et al., 2003; Rescorla & Rosenthal, 2004).

Findings have been equivocal in terms of growth patterns of untimed computational accuracy and general math achievement, with some studies finding evidence for a cumulative pattern in the early elementary grades (Aunola et al., 2004; Chong & Siegel, 2008; McClelland et al., 2006; Morgan et al., 2009; Morgan et al., 2011; Salaschek, Zeuch, & Souvignier, 2014; Stevens, Schulte, Elliott, Nese, & Tindal, 2015), and others finding evidence for a compensatory model (Han, 2008; Mok et al., 2015; Ready, 2013). Still others have noted parallel trajectories; for example, in a longitudinal study examining general math growth from Grades 3 through 10, Rescorla and Rosenthal (2004) found no differences in terms of growth rates between students with a higher initial status in math achievement and cognitive ability compared to lower achieving peers. Further, some authors have highlighted sample heterogeneity in terms of growth patterns (Jordan, Mulhern, & Wylie, 2009; Jordan et al., 2006), and note differences in terms of growth patterns dependent on the subcomponent of math assessed (Salaschek et al.,

2014). For example, in a study investigating the developmental dynamics of different math competencies across Grade 1, Salascheck et al. (2014) found that for untimed calculation, the majority of children followed a compensatory trajectory, although a smaller portion of the students experienced compensatory growth.

In terms of fluency specifically, across Grades 1 through 5, Geary and colleagues (Geary et al., 2012) found compensatory effects for procedural competence, but cumulative effects in terms of direct retrieval and decomposition skill between groups of typically achieving children and those with math learning disabilities (or low achievement, although the effect was less pronounced). Jordan, Hanich, & Kaplan (2003) found that children with low fluency experienced little growth over time, whereas those classified in the stronger fluency group showed linear growth; although this specific issue was not the focus of their study, this pattern would suggest cumulative growth. Other studies mentioned above (math fluency section) found no relationship between level and slope between groupings of typically achieving children and those with math difficulties (Chong & Siegel, 2008; Jordan et al., 2003). For example, Chong and Siegel (2008) found a compensatory pattern for untimed calculation (procedural), but stability for weakness in math fluency.

Although there is no agreed upon statistical test to explore cumulative and compensatory models, the primary statistical indicator for a cumulative effect is a positive correlation between the intercepts of children's growth trajectories and their slopes (Huang, Moon, & Boren, 2014; Protopapas, Sideridis, Mouzaki, & Simos, 2011; Scarborough & Parker, 2003). Generally, if the slopes favor children whose level of achievement is initially strong, then differences between children are magnified. Conversely, if steeper slopes are seen in children who demonstrate initially weaker skills, then the differences between students' skill level are minimized or cancelled out. If the slopes are the same despite varying initial skill level, the differences between student skill levels persist unchanged. Although these patterns may be seen from a

group perspective by examining mean intercepts and slopes (e.g., comparing groups of children with weaker math skills to those who have stronger abilities), considering that children may have different initial abilities (i.e., intercepts), their slopes (speed at which they gain skill) may also vary. Taking a multi-level model approach such as the one utilized in the current study, allows one to capture such inter-individual variability in a sample of students (i.e., each student is allowed to have their own intercept and slope). In this case, the *variance* in intercepts and slopes can be correlated positively (cumulative), negatively (compensatory), or not at all. In the scenario where variance in growth rates is correlated with the intercepts, data would take on a fan pattern (whether fanning out in the case of the cumulative growth, or fanning inward in the case of compensatory growth). Conversely, in the case where there is zero correlation and zero variance in growth rate, children's growth trajectories will be parallel but offset. In terms of the metric used to detect these effects, previous studies have argued for the use of raw scores over standard scores (Bast & Reitsma, 1998; Stanovich, 1986). These authors highlight the fact that raw scores at different ages are transformed to a distribution with the same variance; therefore, the increase in variance expected over time is lost to the standardization process, effectively masking any cumulative or compensatory effect (Bast & Reistma, 1998). The multi-level modeling strategy used in the current study allows for the examination of the development of cumulative versus compensatory growth patterns over time (see next section).

Having an understanding of factors that are associated with increasing achievement gaps, or those that can support the closure of these gaps could theoretically lead to interventions that improve student prospects, both from an educational and societal point of view (Rivera-Batiz, 1992).

**Measurement and Analysis of Developmental Trajectories in Math Fluency**

The following section summarizes key measurement considerations relevant to longitudinal research on math fluency and the current dissertation. Specifically, potential differences in fluency development according to operation type, curriculum-based measurement as a measure of fluency, and analysis of growth trajectories using Hierarchical Linear Modeling are discussed.

**Potential differences between operation types.** One issue that relates to the measurement of math fluency is the type of operation (i.e., addition, subtraction, multiplication, division) assessed. Few behavioral studies have specifically examined differences in operation types (Barrouillet & Lepine, 2005; Barrouillet et al., 2008; Martens et al., 2011; Robinson, 2001; Siegler, 1987). However, converging evidence from neuropsychological and behavioral studies provide a rationale for considering these abilities separately.

One line of evidence highlights that the strategy selected to solve the basic arithmetic problems may vary according to operation type (Barrouillet & Lepine, 2005; Barrouillet et al., 2008). From a theoretical point of view, the work of Dehaene and colleagues (2003), suggest that operations may be dependent on differing neural codes, which are in turn linked to distinct neural pathways. Namely, whereas simple overlearned calculations (e.g., single digit addition and multiplication) are solved through a direct verbal retrieval route (left-hemisphere cortico-subcortical loop), larger addition problems and subtraction are solved through an indirect route, where operands are coded as representations of quantity (left and right hemisphere inferior parietal areas). Evidence for this theory includes findings that operation skills are dissociable in adults (Dehaene & Cohen, 1997; Dehaene et al., 2003; McCloskey, 1992), where subtraction and multiplication show the clearest differentiation (Barrouillet & Thevenot, 2013; LeFevre et al., 1996; Prado, Mutreja, & Booth, 2014). Similar findings have been seen in the limited

neurofunctional studies involving children (Berteletti & Booth, 2015; De Smedt et al., 2011),

with this dissociation increasing with age (Prado et al., 2014).

From a behavioral perspective, two studies systemically investigated strategies that

children use to solve arithmetic problems of addition (Barrouillet & Lepine, 2005) and

subtraction (Barrouillet et al., 2008) and the relationship between strategy use and working

memory capacity. Collectively, these studies found that third grade children used direct retrieval

far less frequently when solving subtraction (19%) than when solving addition (65%), for the

same problem stems (inverse problems). Further, these authors identified differential

relationships between working memory and operation types. In the case of addition, high

working memory capacity was associated with faster problem solving (including small sums),

and children with stronger working memory were less reliant on algorithmic procedures

(Barrouillet & Lepine, 2005). In contrast, in a follow-up study, Barrouillet and colleagues (2008)

found no association between working memory and the speed of direct retrieval in subtraction,

although working memory capacity was associated with faster problem-solving though

algorithmic strategies. These results suggest potential differences between the mechanisms with

which addition and subtraction problems are solved fluently (Dehaene, 1992). Conversely, from

a longitudinal perspective, Van de Weijer-Bergsma and colleagues (2015) did not find

differentiation between operation types in terms of their respective association with working

memory.

Further, in a study examining fluency development among 6- to 15-year-olds, Martens et

al. (2011) found that the ratio of correct to incorrect responses over the years indicated

developmental differences according to operation type, such that addition > subtraction >

multiplication > division. As acknowledged by the authors the fact that multiplication and

division are introduced at least one year later than addition and subtraction may in part explain

this difference. However, the difference between addition and subtraction fluency is notable given that these operations are taught simultaneously.

Many previous studies that have included measures of fluency have used mixed operation measures (Chong & Siegel, 2008; Petrill et al., 2012), which may mask differences linked to operation type. Further, of the studies that have considered single measures, most have focused on addition (e.g., Barrouillet & Lepine, 2005; Schutte et al., 2015; Vasilyeva et al., 2015), and research including measures of subtraction fluency is rarer (Barrouillet et al., 2008; Robinson, 2001). Nonetheless, the aforementioned studies provide a rationale for considering operation types separately. Thus, in this dissertation, separate measures are used to assess addition and subtraction fact fluency. A mixed measure (i.e., addition, subtraction, and multiplication) is also used as a point of comparison.

**Mathematics curriculum-based measurement**. Curriculum-based measurement (CBM) is a standardized procedure used to assess the level and trend of specific skill areas (Deno, 1985; Hosp, Hosp, & Howell, 2007). Traditionally, this measurement tool is primarily used in educational settings to screen for ability, monitor progress, make determinations regarding programming and placement, and to evaluate the effectiveness of programs (Deno, 2003). In the area of mathematics (mathematics curriculum-based measurement, M-CBM), single skill measures (referred to as "probes") are designed to assess a narrow range of ability (e.g., addition, subtraction, multiplication, division) (Fuchs & Deno, 1991). These are also referred to as "robust indicators", as they aim to tap into a core competency that is not necessarily reflective of the curriculum, but geared toward assessment of progress toward skill mastery (Christ, Scullin, Tolbize, & Jiban, 2008; Foegen, Jiban, & Deno, 2007). These single skill probes are in contrast to General Outcome Measures (GOMs) which are designed to capture content relevant to a given

grade curriculum (end of year) (Christ & Vining, 2006). Single skill probes are used in the current study.

A significant difference between CBM probes and commonly used standardized assessments of fluency seen in the literature (e.g., Woodcock-Johnson Math Fluency; Woodcock, McGrew, Mather, & Schrank, 2001) lies in the scoring. In addition to a timed component, M-CBM scores allow for partial credit, as scores represent correct digits (in the correct placement) per minute (CDPM), rather than the number of correct responses per minute. This makes the measure particularly sensitive to improvement, as illustrated by a simple example. Consider, 11 + 7 = 18: A child who correctly responds 18 would receive credit for two digits correct, a child that responds with one error (e.g., 19) would receive credit for the correct "1" value correctly placed in the tens position, whereas a child who responds incorrectly (e.g., 20) would not receive credit. A child who can obtain 40 digits correct per minute using this scoring would arguably have stronger skill than a child who can calculate these 40 digits, but requires double the time to do so.

Research on the psychometric properties of M-CBM has shown that single and multiple skill CBM probes measure slightly different, though related constructs (Foegen et al., 2007; Hintze, Christ, & Keller, 2002). Studies on single skill probes have found adequate to good levels of alternate form reliability (.73 for Grade 3, .93 for Grade 4 (Espin et al. (1989) as cited in Foegen et al. (2007) and .92 for Grade 4 as per Thurber et al. (2002)). Regarding criterion validity, correlations between single skill probes and other math measures (e.g., computation, word problems) ranged from .30 to .60 (Foegen et al., 2007); moderate correlations are expected given that the narrow skill of math fluency is a separate construct from general math measures that tap multiple skills. Single skills probes have been found to have very little variance in difficulty (i.e., aside from the digits used, the algorithms remain constant) across parallel probes; they are homogeneous in terms of content, stimuli, and the requisite skill set assessed (Hintze et

al., 2002). Hintze and colleagues (2002) demonstrated that a 2-minute single skill (addition, subtraction, multiplication, and division) assessment provided exceptionally high levels of dependability from a single administration (i.e., greater than .95). Further, these authors found that upward of 80% of the measurement variance was accounted for by individual and developmental (age) variance among participants, indicating very little unsystematic random error. The conclusion of this study was that single skill M-CBM is of sufficient quality to make both criterion- and norm-referenced decisions (Hintze et al., 2002). This is relevant in light of the fact that longitudinal models require measurement invariability, or measurement of the same construct over time (Singer & Willett, 2003). Conversely, multiple-skill M-CBMs are more heterogeneous, and are more susceptible to differences between parallel assessment forms (see Christ & Vining, 2006; Hintze et al., 2002; Methe, Briesch, & Hulac, 2015).

M-CBM measures have been used in developmental math research as a measure of fluency (Allinder et al., 1992; Allinder & Eicher, 1994; Fuchs et al., 1993; Fuchs et al., 2005; Fuchs et al., 2006; Graney et al., 2009), albeit to a lesser degree than norm- or criterion-referenced tests. Unlike most standardized measures of achievement which are designed to assess a student's standing in comparison to same-age or same-grade peers at a yearly interval or greater, CBM were designed to assess individual growth in skill at shorter testing intervals (Shin, Espin, Deno, & McConnell, 2004). Further, as highlighted by Shin et al. (2004), measurement error associated with growth estimates decreases with increasing data points, leading to higher reliability of growth parameters. The argument for the use of a sensitive measure of skill growth has been made in relation to the summer learning loss (Patton & Reschly, 2013) since measures assessing skill loss over the summer must be sensitive to changes in a relatively short period of time.

**Hierarchical linear modeling in the assessment of growth**. In the current dissertation, analysis of longitudinal data is performed using growth-curve modeling. Growth curve modeling techniques are useful when the question of interest involves examining within-person (i.e., intra-individual) change in a given phenomenon over time, as well the between-person (i.e., inter-individual) differences in these developmental trajectories. For example, throughout the early elementary years, children's math fluency skill increases over time (within-person change), but some individuals may gain skills more rapidly than others (between-person difference). Hierarchical linear modeling (HLM) is a statistical approach used to conduct growth curve modeling, and is an extension of regression analysis for data that are "nested", or hierarchically organized (see Raudenbush & Bryk (2002), Bickel (2007), Little et al. (2009), and Woltman, Feldstain, MacKay, & Rocchi (2012) for more detailed overviews, though portions relevant to the current dissertation are briefly reviewed here). This allows for the analysis of longitudinal data that would otherwise violate the assumption of independence in regression. In the case of longitudinal data, repeated measures are seen as "nested within" individuals, as they are likely to be more closely related to each other than to measures drawn from another individual.

Hierarchical linear model data is structured in at least two levels. For longitudinal data, Level-1 represents the repeated measurements (time), which are "nested" within Level-2, the individual. Therefore, Level-1 captures within-person change over time (e.g., captured by the mean level and growth rate), while Level-2 allows for the estimation of between-person variability (e.g., how much individuals differ in terms of their trajectories).

As an example, the basic Level-1 model for repeated measures is:

$$Y_{ij} = \beta_{0j} + \beta_{1j} time_{ij} + \varepsilon_{ij} \tag{1}$$

Where $Y_{ij}$, is the outcome measure (e.g., math fluency), for an individual $j$ at time $i$. $\beta_{0j}$ is the intercept, $\beta_{1j}$ is the linear slope, and $\varepsilon_{ij}$ is the residual term.

In contrast to ordinary regression, in HLM, the intercepts and slopes are allowed to vary (modeled as "random effects"), such that each individual has their own growth curve (i.e., their own intercept and slope). Although individual coefficients are not estimated, between-person differences are gleaned from estimates of means for the group(s) ("fixed effects"), variances ("random effects", how the means vary), and covariances (whether and how means vary together). This creates the Level-2 equations, or the "slopes-as-outcomes" equations:

$$\beta_{oj} = \gamma_{00} + u_{0j} \qquad\qquad (2)$$
$$\beta_{1j} = \gamma_{10} + u_{1j}$$

$\gamma_{00}$ and $\gamma_{10}$ are the means of the distributions of the Level-1 (the individual children's) intercepts ($\beta_{0j}$) and slopes ($\beta_{1j}$) respectively. The terms $u_{0j}$ and $u_{1j}$ are the Level-2 residuals which represent the deviation (spread around) the mean, $\gamma_{00}$, and slope , $\gamma_{10}$, respectively; $u_{0j}$ and $u_{1j}$ are assumed to be normally distributed with means of 0, variances of $\tau_{00}$ and $\tau_{11}$, and covariance of $\tau_{10}$. This covariance provides a sense of whether the between-person variance in the intercept and variance in slope change together across individuals, thus providing a measure of whether development follows a cumulative or compensatory pattern over time. Further, HLM allows for the meaningful partitioning of the variance (e.g., variance in the overall levels of children's growth trajectories is separate from the variance in slopes), thus allowing for in-depth examination of developmental phenomena across a distribution (Little et al., 2009).

This basic model described above can be expanded at Level-1 to include more complex growth terms (e.g., quadratic, cubic, within-grade versus between grade terms), as is discussed further in Chapter II. Overall, this approach allows researchers to answer questions relevant to the fixed part of the model regarding the functional form of the developmental trajectory (e.g., flat, linear, quadratic), as well as questions relating to the relationship between individual growth curves (e.g., Do individuals vary in terms of their fluency levels and rates of growth?).

The basic model can also be expanded to include time-invariant predictors at Level-2. Adding predictors to the model provides insight into relationships between constructs and the overall mean levels or slopes of the individual growth curve (i.e., fixed portion). This creates implied cross-level interactions, which can be interpreted as Level-2 predictors moderating the effect of time (development) on the outcome (Bickel, 2007; Little et al., 2009). Taking working memory as an example, we can test a hypothesis about how working memory relates to the development of math fluency, that is, how does it relate to the levels (means) of fluency, and how does it relate to the rate of skill growth (linear, quadratic) over time. An estimate of this effect is obtained by examining the fixed portion of the model. The random effects portion of the model gives a sense of the spread (variance) in the levels of curves and/or slope distribution; by comparing the random portions of the model before and after the addition of working memory, it is possible to estimate how much of the between-student variance is explained when working memory is taken into account.

HLM offers some distinct advantages to older approaches to analyzing longitudinal data, such as repeated measures ANOVA (see Shin et al., 2004, for a review). The first benefit is the fact that data can be collected at varying time points, which is important in research considering the effects of the summer since the school year interval is longer than the summer month one (although see Chapter II, below). Second, HLM efficiently handles missing data, and in fact allows for missing data at all but the highest level of the model (i.e., Level-2 in the current study). This is a significant benefit in longitudinal data as it minimizes data loss (e.g., if a child was absent for one testing day, but present for the three other administration days, those data can be used). Further, HLM also permits the modeling of "planned missingness", which is the case of an accelerated longitudinal design that is used in the current dissertation, as explained in Chapter II, below. Notably, many of the studies examining the effects of the summer on math ability

have utilized ANOVA methods, comparing pre- and post-break measures from a categorical perspective (Allinder et al., 1992; Allinder & Eicher, 1994), although more recent models have adopted a longitudinal approach (Alexander et al., 2001; Davies & Aurini, 2013; Downey et al., 2004; Vale et al., 2013). Further, HLM has been specifically highlighted as useful in studies analyzing CBM data (Shin et al., 2004).

**Potential Determinants of Math Fluency Development**

In addition to describing growth patterns, an important contribution of longitudinal studies is the exploration of potential factors that affect development in a given domain. Previous studies that have modeled the effects of the summer have considered how personal traits affect losses over the summer months, although variables have been largely constrained to sex, SES, grade, and IQ (Cooper et al., 1996; Downey et al., 2004). Studies in cognitive and developmental psychology, on the other hand, have evaluated the effects of a number of different cognitive and behavioral predictors, such as IQ (Geary, 2011a), working memory (Geary, 2011a; LeFevre, Berrigan et al., 2013), phonological processing (Barnes et al., 2014; Fuchs et al., 2006; Hecht et al., 2001), processing speed (Bull & Johnston, 1997), number sense (Cowan et al., 2011; LeFevre et al., 2010; Salaschek et al., 2014), and classroom inattention (Fuchs et al., 2006) on the development of math. Although these two frameworks have largely remained separate, it is conceivable that individual differences in cognitive abilities would not only be associated with development over time, but that these patterns may interact with the seasonal effects of the summer. Therefore, the current study investigates the effects of two domain-general abilities, working memory and classroom attention (while controlling for sex and parent level of education), which have been the focus of a number of prior studies in terms their respective roles in the development of general math achievement, computational accuracy, and word problems

(Giannopulu, Escolano, Cusin, Citeau, & Dellatolas, 2008; Holmberg & Bolte, 2014; Pingault et al., 2014; Raghubar, Barnes, & Hecht, 2010).

**Sex.** Sex differences in mathematical ability tend not to be significant when looking at math from a general perspective (Aunola et al., 2004; Lachance & Mazzocco, 2006; Lindberg, Hyde, Petersen, & Linn, 2010). However, previous research in the realm of direct retrieval has consistently found that boys tend to use retrieval more frequently than girls (Bailey et al., 2012; Carr & Jessup, 1997; Carr & Davis, 2001; Laski et al., 2013), which may lead to sex-based differences in performance on tasks including arithmetic (Royer et al., 1999). For example, Carr and Jessup (1997) found that when solving simple addition and subtraction problems, girls were more likely to favor the use of slower overt strategies (e.g., finger counting), whereas boys showed a greater tendency to use a direct retrieval strategy from memory. This effect was found to be significant by the end of Grade 1, despite no sex differences at the beginning of first grade. Notably, these authors found that there were no sex differences in terms of the number of total correct responses (i.e., accuracy) (also see Carr & Davis, 2001), which as proposed by the authors, may explain non-significant findings regarding measures of accuracy (Imbo & Vandierendonck, 2007; Lachance & Mazzocco, 2006). In a longitudinal study investigating preference of using retrieval across Grade 1 to 6, Bailey and colleagues (2012) found that in all six grades, boys showed a preference for retrieval over other potential strategies more frequently than girls, and regardless of retrieval accuracy. Notably, depending on the grade studied, there was not always a sex-based difference in terms of accuracy (e.g., no difference in terms of accuracy in Grade 1, and girls being more accurate than boys in Grade 2). However, boys' retrieval accuracy increased across the grades, and boys outperformed girls in terms of accuracy by sixth grade. Further, the preference for retrieval was not due to stronger working memory skills, as there were no sex differences on these measures.

As mentioned above, considering studies that have reported neurophysiologically distinct representations of operation types (Dehaene et al., 2003), sex differences may also vary as a function of operation type. Indeed, Martens and colleagues (2011) that when looking specifically at fluency, sex differences emerged favoring boys for addition, and to a lesser extent, subtraction (though not multiplication or division). However, these differences were primarily seen in older children, a trend that was also noted by Royer et al. (1999). Other authors have not noted sex differences (Jordan et al., 2003b), highlighting the fact that this issue remains equivocal with respect to fluency.

**Parent level of education.**  Parent level of education has been shown to be an indicator of socio-economic status (SES), that is strongly associated with children's academic achievement, both in the early (Melhuish et al., 2008) and middle childhood (Davis-Kean, 2005) years (Davis-Kean, 2005; Entwisle & Alexander, 1990; Melhuish et al., 2008). Using data from a national cross-sectional sample, Davis-Keane (2005) found that parent education (highest education level in the household regardless of which parent), exerted both direct and indirect effects on children's reading and math achievement for European-American families. For African-American families, parent level of education exerted significant indirect effects on achievement, through parents' expectations of their children's educational outcomes, parenting practice of reading to children, and parental warmth in interactions. Notably, Davis-Keane found that parent level of education was a significantly stronger predictor than income on child achievement. Her study highlighted how parent education can influence both the home environment as well as the manner in which parents interact with their child to promote academic success. Parent education has also been found to have an impact on summer learning loss for arithmetic (Paechter et al., 2015). Further, as mentioned previously, children's SES has clearly been linked to differential effects of the summer, primarily in reading (Cooper et al.,

1996; Davies & Aurini, 2013), but also in general mathematics (Downey et al., 2004). Therefore, in the current dissertation, parental education is included and serves as an indicator of SES.

**Working memory.** Working memory is a limited-capacity, multi-component cognitive system that allows for the maintenance and manipulation of information "on-line" for a short period of time (Baddeley, 1992; Baddeley, 1996; Miyake & Shah, 1999). Much of the developmental research involving math development and working memory has been conducted within the framework of Baddeley's tripartite model, which is conceptualized as three interdependent systems that differ between stimulus modality (auditory-verbal versus visual-spatial) and processing requirements (storage only versus storage plus manipulation) (Baddeley, 1996). In this model, the central executive is a domain-general attentional controller that commands various functions including planning, sequencing, and monitoring of information in active storage in order perform complex cognitive tasks. The phonological loop and visual-spatial sketchpad are responsible for the storage and rehearsal of speech-based and visual-spatial information, respectively. However, other lines of evidence demonstrate that storage and manipulation components are closely related constructs, and show significant overlap despite specific sources of variance (e.g., Colom, Shih, Flores-Mendoza, & Quiroga, 2006; Engle, Tuholski, Laughlin, & Conway, 1999). The overlap appears to be strongest in the visual-spatial domain (Alloway, Gathercole, & Pickering, 2006; Metcalfe, Ashkenazi, Rosenberg-Lee, & Menon, 2013; Miyake, Friedman, Rettinger, Shah, & Hegarty, 2001), where some authors have found short-term memory and working memory to be indistinguishable (e.g., Colom et al., 2006; Miyake et al., 2001). Specific to mathematics, the executive portion of working memory has been strongly linked to math ability in general (Friso-van den Bos, van der Ven, Kroesbergen, & van Luit, 2013; Peng, Namkung, Barnes, & Sun, 2016; Raghubar et al., 2010), although its specific involvement in basic math fluency development has been somewhat equivocal.

Working memory is an important component of Geary's (1993) influential theory on the nature of math disabilities, which posits that deficits in this domain could lead to a failure in consolidating math facts to long-term memory when problems are not successfully paired with their answers before working memory decays (Geary et al., 1991; Lemaire & Siegler, 1995), making these facts less available for direct retrieval. This could the case for children who have more limited working memory capacity, or if the use of slow or ineffective strategies tax working memory resources (Geary, 1993; Imbo & Vandierendonk, 2007). From a cross-sectional perspective, as mentioned above, Barrouillet and Lepine (2005) found that third and fourth-grade students with stronger working memory skills used direct retrieval more frequently when solving simple addition problems compared to their peers with weaker working memory skills. These authors postulated that working memory capacity influences the formation of associations between problem stems and their answers (achieved by counting), whereby working memory inhibits the activation of associations between problems stems and incorrect (though related) responses, thus leading to a greater resistance to interference and enhanced ability to retrieve correct facts directly from memory (Barrouillet & Lepine, 2005). In their follow-up study, Barrouillet and colleagues (2008) found that stronger working memory also predicted faster problem solving of subtraction facts using an algorithmic procedure, although unlike addition facts, the rate and frequency of direct retrieval of subtraction facts were not significantly related to working memory capacity. Work by Imbo and Vandierendonk (2007) also highlighted the influence of working memory capacity and retrieval skills for simple addition problems in 10 to 12-year-olds. Using a dual-task procedure to increase the working memory load (functionally lowering working memory capacity), these authors found that taxing working memory led to slower direct retrieval times as well as slower execution of other strategies such as decomposition and counting. However, the authors also found that increasing working memory load exerted less of an effect with advancing age, in parallel with the use of more efficient skills

(e.g., when retrieval was used more frequently and when other procedural strategies were used with greater efficiency). Further, more frequent retrieval and efficient counting strategies reduced working memory requirements.

Longitudinal studies have also identified a relationship between working memory capacity and fact fluency in the elementary grades (LeFevre et al., 2013; Martin, Cirino, Sharp, & Barnes, 2014), including its role as a predictor of math fluency growth over time (LeFevre et al., 2013). This relationship may be linked to the finding by Geary and colleagues (Geary et al., 2012) that stronger central executive skills (as indexed by both verbal and verbal-numerical tasks) predicted maturation of children's (Grades 1-5) counting strategy use, from a *min* counting strategy (e.g., $4 + 5 = 5 + 1 + 1 + 1 + 1 = 9$) to a more sophisticated decomposition strategy (e.g., $4 = 2 + 2$, therefore $4 + 5 = 5 + 2 + 2 = 9$). Intact counting abilities would contribute to math fluency development as they result in representations of basic facts in long-term memory (Siegler & Shrager, 1984), and because they are often used as a back-up procedure when direct retrieval fails (Geary et al., 1991; Geary, 1993; Geary et al., 2012; McKenzie, Bull, & Gray, 2003). It would follow that a more rapid transition to greater counting sophistication would lead to more efficient strategy use, and therefore to faster development of arithmetic fluency. Further, children with working memory weaknesses may rely more heavily on finger counting and may commit more counting errors than those with stronger working memory skills, which would also impede fluent arithmetic (Geary, Hoard, Byrd-Craven, & DeSoto, 2004).

Conversely, other studies did not find working memory to be a predictor of math fluency skill (Vanbinst, Ceulemans, Ghesquiere, & De Smedt, 2015), or found that it was no longer a unique predictor of fluency when other factors, such as classroom inattention were taken into account (Fuchs et al., 2005; Fuchs et al., 2006). As argued by Geary et al. (2012), a lack of significant findings may be in part related to the age at which working memory plays a role in arithmetic, such that working memory may take on greater importance during the earlier grades

while skills are being acquired (e.g., reliance on more memory-taxing strategies such as decomposition) (Peng et al., 2016; Raghubar et al., 2010), with decreasing importance in advancing grades, as the strategy mix moves toward a greater emphasis on direct retrieval compared to procedural strategies (Ackerman, 1988; Geary et al., 2004; Imbo & Vandierendonck, 2007). Conflicting results may also be in part related to the type of working memory measure used. Specifically, in a recent meta-analysis (Peng & Fuchs, 2014) highlighted the role of domain specificity in working memory tasks, with verbal-numerical working memory tasks (e.g., backward digit span) emerging as a stronger predictor of mathematical ability (Martin et al., 2014) compared to a strictly verbal (e.g., sentence span) ones (e.g., Fuchs et al., 2005; Fuchs et al., 2006).

Although not as frequently used as a measure of working memory compared to verbal working memory measures, visual-spatial working memory may be a particularly important predictor of basic arithmetic abilities (Ashkenazi, Rosenberg-Lee, Metcalfe, Swigart, & Menon, 2013; Lukowski et al., 2014; Metcalfe et al., 2013; Raghubar et al., 2010), particularly in the elementary grades while the skill is still being acquired (Holmes & Adams, 2006).  For example, recent neuroimaging studies identified visual-spatial working memory as being specifically linked to arithmetic problem solving in children aged 7-9 years, whereas verbal working memory was not (Ashkenazi et al., 2013; Metcalfe et al., 2013). Visual-spatial working memory may support fluency development through the formation of mental models (representation of quantity mapped onto objects and their movements), which young children use when solving nonverbal addition and subtraction problems (Bisanz, Sherman, Rasmussen, & Ho, 2005; Huttenlocher, Jordan, & Levine, 1994; Rasmussen & Bisanz, 2005).  Visual-spatial working memory has also been associated with the representation of quantity along an internal mental number line (Berteletti, Man, & Booth, 2015; Geary et al., 2007; Rotzer et al., 2009), which children may use to solve basic arithmetic problems that are not automatically retrieved (Gunderson, Ramirez,

Beilock, & Levine, 2012). Visual-spatial working memory may also play a role in fluency development by supporting the maturation of counting strategies (Geary, Hoard, Byrd-Craven, Nugent, & Numtee, 2007). Finally, visual attention (visual working memory) is a key component in LeFevre and colleagues' Pathways to Mathematics model, acting as a longitudinal predictor of the development of various branches of math ability, including numeration, number line knowledge, and untimed calculation (LeFevre et al., 2010). Recent work using this model found that the working memory pathway (refined to also index the central executive and phonological loop) was a unique predictor of math fluency ability in students in Grades 2 and 3, when controlling for sex, parent level of education, and processing speed (Sowinski et al., 2015).

Further, as previously mentioned, the importance of working memory in fluency development may be dependent on operation type (Barrouillet & Lepine, 2005; Barrouillet et al., 2008). Finally, it has also been suggested that the relative effect of verbal and visual-spatial working memory on arithmetic may shift in accordance with development and experience, with basic arithmetic skills being contingent on the visual-spatial skills during the acquisition phase (Holmes & Adams, 2006; Raghubar et al., 2010), and verbal skills becoming more important in supporting the direct retrieval of consolidated facts that are based on a verbal code (Dehaene et al., 2003; Holmes & Adams, 2006; McKenzie et al., 2003). This assertion is supported by a recent study by Van de Weijer-Bergsma and colleagues (2015), who found that the relationship between visual-spatial memory and math fluency was stronger in the early elementary grades, and weakened over time. Conversely, they found that the link between verbal working memory and math fluency grew stronger with increasing grades. Therefore, despite the clear importance of working memory math ability in general, there is still no consensus on its specific role (whether assessed by verbal or visual-spatial means) in the development of math fluency. Therefore, measures of verbal working memory and visual-spatial working memory are included in the current study to examine this issue further.

**Behavioral inattention**. Another domain-general predictor of mathematical achievement that has received increased attention in recent years are behavioral ratings of inattention, particularly teaching ratings of children's inattention in the classroom. Inattention has been found to be predictive of academic underachievement, including weaknesses in math abilities (Polderman, Boomsma, Bartels, Verhulst, & Huizink, 2010).

Previous studies have identified a robust finding that the dimension of inattention is associated with poor math achievement, both from cross sectional (Rodriguez et al., 2007) and longitudinal perspectives (Currie & Stabile, 2006; Massetti et al., 2008; Merrell & Tymms, 2001; Pingault et al., 2011). This relationship has been identified using measures of general math ability (Fitzpatrick & Pagani, 2013; Garner et al., 2013; Gold et al., 2013; Holmberg & Bolte, 2014; Massetti et al., 2008), and specific sub-skills, including math fluency (Fuchs et al., 2005; Fuchs et al., 2006; Lewandowski, Lovett, Parolin, Gordon, & Codding, 2007; Martin et al., 2014), computation in terms of accuracy (Fuchs et al., 2006; Li & Geary, 2013; Raghubar et al., 2009), and word problems (Fuchs et al., 2006; Swanson, 2011). Results of a recent longitudinal study (spanning Kindergarten to Grade 5) suggest that children who demonstrate symptoms of classroom inattention earlier (i.e., beginning in Grade 1) and those who demonstrate these symptoms persistently over two years (Grades 1 and 2) were at greater risk of persistent math (untimed calculation) and reading weakness compared to children with no concerns regarding attention skills or whose inattention was only noted later (i.e., Grade 2) (Rabiner, Carrig, & Dodge, 2016). The dimension of inattention, specifically, has consistently been shown to be a stronger predictor of mathematical ability compared to the dimension of hyperactivity for children diagnosed with ADHD (Lee & Hinshaw, 2006; Massetti et al., 2008). This distinction is also seen in community samples (Garner et al., 2013; Pingault et al., 2014). Further, ratings of inattention as provided by teachers are stronger predictors of achievement than parent ratings of inattention (Garner et al., 2013).

With respect to math fluency specifically, children with ADHD experience persistent difficulties in this domain (Ackerman et al., 1986; Lewandowski et al., 2007; Zentall, 1990; Zentall & Smith, 1993). Although computational accuracy differentiates students with ADHD from their typically achieving peers throughout elementary school (Mariani & Barkley, 1997; Zentall, Smith, Lee, & Wieczorek, 1994), by middle school, only math fluency continues to distinguish these two groups (Zentall, 1990). Lewandowski et al. (2007) found that children diagnosed with ADHD were weaker on math fluency tasks than their typically achieving peers, and that this difference was not fully accounted for by processing speed ability. Further, in individuals with ADHD, difficulties with math fluency persist into adulthood (Biederman et al., 2005).

Symptoms of ADHD occur on a continuum in the population (Polderman et al., 2007) and the association between inattention and math fluency extends to community samples, although fewer studies have looked at this relationship specifically. For example, Fuchs and colleagues (2005) found that teacher-rated inattention was a unique predictor of math basic math fluency skills in Grade 1 students, and that other cognitive factors such as processing speed and working memory (as measured by digit span) were not significant once inattention was taken into account. Further, Fuchs et al. (2006) found that teacher-rated inattention and processing speed were significant predictors of algorithmic fluency, although working memory was not. Further, in terms of strategy choice, Geary et al. (2012) found that classroom inattention predicted the rate at which children transitioned from less efficient procedural strategies to more sophisticated memory-based strategies when solving basic arithmetic. Conversely, other studies have not found inattention to be a unique predictor of fact fluency when other predictors such as verbal and visual-spatial working memory, phonological processing, and number processing abilities were accounted for (Martin et al., 2014). Accordingly, teacher-rated classroom inattention is included as a predictor variable in the current thesis in order to examine this issue

43

further.

        **Relationship between working memory, inattention, and mathematics.** Empirical research highlights a close relationship between working memory, attention, and mathematics (Alloway, Gathercole, Kirkwood, & Elliott, 2009; Alloway, Elliott, & Place, 2010; Gray, Rogers, Martinussen, & Tannock, 2015; Rogers, Hwang, Toplak, Weiss, & Tannock, 2011). First, among children with ADHD, the dimension of inattention, and not hyperactivity/impulsivity, appears to be related to both verbal and visual-spatial working memory abilities, regardless of age, verbal intelligence, reading, and language skill (Martinussen & Tannock, 2006). Among non-referred children, those with weak working memory tend to perform poorly on math tasks, regardless of IQ, and these children also tend to be rated by their classroom teachers as being inattentive (Alloway et al., 2009). In a study on 5 to 7-year-old children, Fuchs and colleagues (2010) found that classroom attention was predictive of math ability (basic fact fluency and word problems) over and above measures of number sense, and was moderately correlated with measures of central executive component of working memory (although working memory itself did not account for unique variance). Using a longitudinal model, Rennie et al., (2014) found that working memory assessed in Grade 1 was predictive of third-grade mathematical achievement (using a composite of math fluency, arithmetic, and problem-solving) in students displaying high ADHD symptoms, although not in students displaying low symptoms of ADHD. In terms of the nature of the relationship between teacher-rated inattention and working memory in relation to achievement, recent longitudinal research from our lab (using a sample nearly identical to the one presented in the current dissertation) identified visual-spatial working memory as a partial mediator in the predictive relationship between classroom inattention and math fluency (addition and subtraction assessed one year later) for boys in a community sample of students (Gray et al., 2015). Notably, working memory

44

did not mediate the relationship between inattention and reading fluency. Finally, other lines of evidence have proposed that math fact retrieval difficulties may be linked to problems inhibiting irrelevant information from working memory during fact retrieval (Geary et al., 2012; Passolunghi & Siegel, 2004). Therefore, the overlap between attention and working memory underscores the importance of integrating both domains in the study of math development.

**Rationale of the Doctoral Research**

Overall, there has been increasing interest in mathematics research over the past 20 years, and growing focus on longitudinal studies on the development of math ability. However, the developmental dynamics specific to the basic skill of math fluency has received relatively little research attention. Further, longitudinal investigations focusing on fluency development have varied in terms of study length, with many spanning one school year and others spanning a number of years, making the shape of developmental trajectories difficult to ascertain. A fundamental issue with these studies, however, is that they do not take into consideration the structure of the school year, specifically the effects of the summer months, which previous research has demonstrated is associated with slowed growth or skill loss (Cooper et al., 1996). Therefore, research is needed to delineate growth patterns, taking into account both within-grade effects (growth specific to time children spend at school) and between-grade effects (spanning a number of years, including the effects of the summer months). When considering the body of research on summer learning, studies have generally been constrained to the realm of educational psychology and policy. These studies primarily focused on three main areas: a) demonstrating the effect of summer slowing or loss on achievement, b) highlighting the net effect of schooling on achievement, and c) explaining the increasing academic gaps between children from low and high socio-economic backgrounds (Cooper et al., 1996; Verachtert et al., 2009). While prior research has examined the effect of individual differences in sex, socioeconomic status, and IQ

on learning and learning slowing over the summer (Cooper et al., 1996; Davies & Aurini, 2013; Downey et al., 2004), to my knowledge studies to date have not examined the unique contributions of working memory and classroom inattention on fluency development within a framework which accounts for both within- and between-school year growth. The current study contributes to this literature by attempting to merge this area of educational research with developmental/cognitive research. The goal here is to identify key interactions between seasonal effects (within-year and between-years, including the summer) and individual factors stemming from cognitive research found to be associated with math development. Although I examined community-level data, the identification of these factors may be particularly important flags for later weakness in math, or targets for early intervention. The aforementioned issues are highly relevant to educational and clinical practice, considering that a significant minority (approximately 10 percent) of students experience persistent low achievement in mathematics despite average abilities in other areas, in addition to the approximately 5-8% with mathematical learning disabilities (Geary, 2011b). These estimates suggest that there are a number of children who struggle in math who may be ineligible for special education services. Without knowledge of variability of the growth trajectories of typically developing children, support geared at addressing areas of individual weaknesses may be lacking, leading to missed opportunities for these children to achieve to their potential. This would be an especially important point in the case of growth trajectories that follow a cumulative growth pattern.

Therefore, the overarching goal of the current dissertation is to address these gaps in the literature, attempting to merge and extend educational and developmental/cognitive lines of research, by answering the following questions:

1. What is the best fitting functional form of math fluency (addition and subtraction) when taking into account both between- and within- year growth, and is there evidence of a summer slowdown in math fluency?

2. How much do children's math fluency (addition and subtraction) trajectories vary between students, and does this variability increase (cumulative pattern) or decrease (compensatory pattern) across Grades 1 through 4?

3. What is the predictive value of verbal working memory, visual-spatial working memory, and teacher-rated inattention on math fluency (addition and subtraction) development when controlling for sex and parent level of education?  What are the specific influences of these predictors on the different phases of growth?

4. Are there differences between addition and subtraction growth curves, in terms of the growth curves themselves (e.g., shape), longitudinal growth patterns (cumulative versus compensatory), and their respective relationship to predictor variables outlined in Question 3?

The goals of this study were accomplished through growth curve analysis (HLM) using an accelerated longitudinal design, with a community sample of children who were in Grades 1 to 3 at study entry. The provincial curriculum focuses on explicit teaching of addition and subtraction starting in Grade 1. Children in each cohort were assessed at four time points over the course of two years, which allowed for modeling of higher level functional forms, as well as the ability to "link" overlapping trajectories in order to simulate a single developmental process. Single skill addition and subtraction curriculum-based measurement (CBM), which have sound psychometric properties (Hintze et al., 2002; Thurber et al., 2002) were used as a measure of fluency. Separate analyses were completed for CBM addition and CBM subtraction to examine potential developmental differences between operation types. Further, current theories of

mathematical cognition suggest that skill in different operations are dependent on separate neural networks, and that there is evidence that operation skill is dissociable in both adults (Dehaene & Cohen, 1997; Dehaene et al., 2003; Mccloskey, 1992) and children (Berteletti & Booth, 2015; Prado et al., 2014). It is, therefore, conceivable that that there would be differences in developmental trajectories or with their relationship to associated cognitive/behavioral predictors based on the operation type; this issue has received relatively little attention, and results have been mixed, with some authors finding differences in the contribution of working memory based on operation type (Barrouillet & Lepine, 2005; Barrouillet et al., 2008), although others have not (Van de Weijer-Bergsma et al., 2015).  A third analysis using a mixed standardized math fluency probe was used as a point of comparison.

In accordance with previous research in math development, it was hypothesized that a general curvilinear increase in skill would be seen across the years, but with practice being critical to fluency development (Bailey et al., 2012; M. K. Burns, 2005; Nelson, Burns, Kanive, & Ysseldyke, 2013), it was also expected that skills would be gained over the school year more rapidly than over the summer months (Allinder et al., 1992; Patton & Reschly, 2013).  Further, given that fluency has been found to be a strong predictor of the growth of arithmetic ability and the rate at which children increased their use of cognitive (e.g., counting mentally, retrieval from memory) strategies (Carr & Alexeev, 2011) and that fluency leads to greater retention (Binder, 1996), it was expected that that math fluency development follows a cumulative pattern of development. Based on previous findings, it was also expected that verbal working memory, visual-spatial working memory, and classroom inattention would emerge as significant predictors of math fluency development.

**CHAPTER II: Methods**

**Participants**

Participants (*n* = 204*)* of the current study were a subset of a larger sample (*n* = 524) of elementary school children, aged 6 to 9 years, who had been recruited for a study examining longitudinal relationships between classroom attention, working memory, and academic achievement (Social Sciences and Humanities Research Council of Canada; SSHRC: project # 410-2008-1052). The original sample was recruited from seven public elementary schools (constituting 20% of the 33 schools in the board), in a large Canadian school board, containing both suburban and rural schools and serving a socio-economically diverse student population. Inclusion criteria were that participants were recipients of mainstream classroom education in either English or French (29% were in French immersion program), had no significant sensory or physical impairment, had received written informed parental and teacher consent, and had provided assent. As per the Ontario curriculum, mathematics instruction in the participating school board followed a discovery-based approach (Ross, Hogaboam-Gray, McDougall, & Bruce, 2002).

The subset of participants were 214 elementary school students (as well as their teachers) who were randomly selected from the larger pool of participants based on their teachers' ratings of inattention using the Strengths and Weaknesses in ADHD and Normal Behaviors Scale (SWAN (see description below); Swanson et al., 2004; Swanson et al., 2012). Stratifying for sex, 2-3 students in each class from the top, middle, and bottom rank-ordered scores on the SWAN were selected to undergo more in-depth assessment, which included cognitive measures (e.g., working memory). Parents had provided consent with the knowledge that their child may or may not be selected for the in-depth assessments and thus would take part in either 2 or 4 assessments over the 2-year study period. This subset did not differ significantly from the full sample in terms of participant sex and age. Only 10 of the 214 students were in Grade 4, which was significantly fewer than the students in each of the other grades. Thus, to avoid potential skewing of the

50

results, these students were excluded from analyses. Therefore, the final subgroup for this study consisted of 204 students (102 boys and 102 girls), who were in Grades 1, 2, or 3 at study entry (Year 1, see Table 1). These students were followed for 2 years, and so were reassessed in Grades 2, 3, and 4 at Year 2 of the study. Most children spoke English as their primary language at home (95%), and identified as Caucasian (80%). Responding parents were mostly mothers (92%), and 93% of families had at least one parent (responder or spouse) who had completed at least some post-secondary education. With respect to learning exceptionalities (as defined by the Ontario Ministry of Education), 8.8% did not respond to this question, 75% indicated no exceptionality, and 15% identified exceptionalities as outlined in Table 1[3]. Exceptionalities for the current sample are comparable to the incidence rates in the general population (Learning Disabilities Association of Ontario, 2011; Wilcutt, 2012).

[INSERT Table 1. *Participant demographics at study entry (fall of year 1)*]

**Data-Collection Procedures**

The study was approved by the Ethics Committees of the University, The Hospital for Sick Children, and of the participating school boards. Parents and teachers provided written informed consent and each student provided verbal assent prior to participation in the study. Test administration and scoring was completed by psychologists, research assistants, and psychology graduate students, all trained in psychometric test administration. Participating students completed written math fluency measures (addition and subtraction CBM, WJ-III) in

---

[3] Although ADHD does not in itself fall within its own exceptionality category (can be included in any category, according to a student's learning need), ADHD diagnosis was included in the screening measure because extant research indicates a strong link between the dimension of inattention and academic difficulties (e.g., Holberg and Bolte, 2014; Garner et al. 2014; Rodriguez et al., 2007; Massetti et al., 2008), and is therefore relevant to the research questions.

small grade-groups of 3-6 students in a quiet room (e.g., library, empty classroom) in their schools. The fluency assessment took place at 4 time-points: in the fall (i.e., the first term of the academic year) of year 1 (Time 1; November 2009), the spring of year 1 (Time 2; May 2010), the fall of year 2 (Time 3; November 2010); and the spring of year 2 (Time 4; May 2011). Working memory and math fluency measures were administered individually in the spring of each study year (Time 2 and 4). Teacher ratings of classroom attention were collected in the fall of each study year (Time 1 and 3), within two months of the start of the study year, allowing for time for the teachers to become familiar with their students.

**Instrumentation**

**Demographics**. Information regarding participants' sex and their parents' level of education was provided by parents in a background questionnaire. Parents provided their highest level of education (1=Grade 1-8, 2=Grades 9-11, 3=High School/GED, 4=Some College, 5=College Graduate, 6=Some University, 7= University Graduate, 8=Post-College Degree (e.g., Masters)) from themselves and their spouse if applicable; the highest level of education (for either respondent parent or their spouse) was used in the analyses (Davis-Kean, 2005).

**Addition and subtraction fluency.** Students' growth in math fluency was measured using parallel addition and subtraction curriculum-based measurement math probes (M-CBM) from AIMSweb M-CBM. CBMs were selected as the primary measure of fluency as they are highly sensitive to improvement in fluency skill (Thurber et al., 2002). For each operation type, students in Grades 1-3 were given two minutes to write answers to up to 60 basic addition/subtraction facts ("primary" level forms). In Year 2, participants tested in Grade 4 (i.e., Year 2 for students who entered the study in Grade 3), were given two minutes to write answers to up to 84 basic addition/subtraction facts ("intermediate" level forms). Operations were composed of digits from 0 to 12, with no carrying or borrowing. The two form types (primary

versus intermediate) differed only in the total number of problems (one additional problem per row and one additional row per page) so as to avoid potential ceiling effects in older students. Concurrent validity was calculated between CBM measures and the well-validated Woodcock Johnson Tests of Achievement Math Fluency subtest (Woodcock, McGrew, Mather, & Schrank, 2001) ($r = .88$, $p < .001$ for addition; and $r = .87$, $p < .001$ for subtraction), which was also used in the fluency analyses as a point of comparison (see below).

*CBM scoring.* For both CBM operation types, fluency performance was estimated as being the number of correct digits answered correctly per minute. A "correct digit" was defined as a correct digit in the correct place value. For example, for an addition problem such as $7 + 4$, the correct response of "11" would receive 2 correct digits for the correct digits in the correct places. For problems that were either incorrect or incomplete, credit was awarded for correct digits that were also in the correct place value. Using the same example as above, a response of 10 would receive one correct digit for the "1" in the tens position, but none for the "0". The order of the completion of the probes (i.e., addition or subtraction completed first) was randomized. Please refer to Appendix A for further examples.

**WJ-III math fluency.** To provide a standardized comparison to the math CBM measures, students also completed the Math Fluency subtest on the Woodcock-Johnson Tests of Achievement, Third Edition (WJ-III; Woodcock, McGrew, Mather, & Schrank, 2001). In contrast to CBMs, the WJ-III consists of a mix of operations, progressing from addition and subtraction, to multiplication. Children were given 3 minutes to complete this task. Scoring of the WJ-III was done in the standardized manner (i.e., score was the number of correct responses in obtained in 3 minutes). However, in order to be able to compare the outcomes to CBM measures, this score was also transformed into a score representing the number of correct response per minute (i.e., score divided by 3 minutes).

**Working memory.** Components of working memory were assessed for both verbal and

visual-spatial modalities in the spring of each study year. Verbal WM was assessed using the Digit Span test of the WISC-IV (Wechsler, 2003). Both forward (measure of short-term memory) and backward (measure of executive control, working memory) digit spans were administered; standard scores of the backward measure were used as an index of verbal working memory. Visual-spatial WM was measured using the Finger Windows subtest from the Wide Range Assessment of Memory and Learning (WRAML2) (Adams & Sheslow, 2003). The score is the longest sequence correctly reproduced by the child. Both the standardized forward (short-term memory) subtest and an experimental backward version (working memory; Bedard & Tannock, 2008) were administered; raw scores for the finger windows backward task were used in the current study.

      **Inattention**. For each participating student, classroom teachers completed the Strengths and Weaknesses in ADHD and Normal Behaviors Scale (SWAN; J. Swanson et al., 2004; J. M. Swanson et al., 2012) rating scale in the fall (November) of study Year-1 and Year-2. The SWAN includes DSM-IV criteria for ADHD including symptoms of inattention (Questions 1-9) and hyperactivity/impulsivity (Questions 10-18).  Each item is scored on a 7-point Likert scale from +3 ("Far Below Average") to -3 ("Far Above Average"). Thus, the higher the score, the more inattentive or hyperactive/impulsive the teacher rated the child. Only the inattention subscale was used in the current study, because extant research consistently identified this dimension specifically (as opposed to hyperactivity/impulsivity) as a significant predictor of academic outcomes (Garner et al., 2013; Gau, 2011; Massetti et al., 2008; Pingault et al., 2011; Pingault et al., 2014; Rodriguez et al., 2007). Inattention subscale scores were obtained by calculating mean ratings from Questions 1-9 of the SWAN for each student; these served as the basis for dimensional measures of students' inattentive classroom behavior. Reliability analyses for this subscale were performed using SPSS 19.0 (IMB Corp., 2010). For the current sample, Cronbach's alphas were .98 for the teacher-rated inattention subscale, for both study years.

Considering the possibility that children's level of attention may fluctuate from year to year (Rabiner et al., 2010) and/or that different teachers may rate similar behaviors differently, additional analyses were undertaken to determine whether the use of a composite score of inattention was warranted. The correlation between Year 1 and Year 2 inattention scores was $r =$ 0.75, $n = 194$, $p < .000$. Paired t-tests were performed for each grade pair in each school (e.g., inattention ratings for student cohort in a given school, in Grade 1 compared to Grade 2 ratings). Although this yielded 20 comparison groups with very small sample sizes ($n = 5$ to 14), such paired t-tests are feasible (e.g., acceptable statistical power and Type I error rate not typically exceeding a value of 5%), as demonstrated by de Winter (2013). Paired t-tests were non-significant, with the exception of one pair from the Grade 3 cohort ($p = .045$). Closer examination of the participants in this pair group revealed that a single individual was rated as significantly more attentive (difference of 1 SD) by his teacher in Year 2 as compared to his teacher in Year 1. As the inattention ratings were highly correlated, and the distributions of inattentive behavior did not differ significantly from one year to the next on the whole (19 out of 20 groups), a mean teacher-rated inattention score was used in multilevel analyses.

**Statistical Analysis**

Prior to analyses, all variables were examined for accuracy of data entry, missing values, and fit between their distributions and the assumptions of multivariate analysis using SPSS version 19.0. The distributions of addition and subtraction CBM and WJ-III fluency for each testing point were normal, with the exception of both CBMs in the fall of Year 1, where the distributions were positively skewed (skew = .61 $SE = .17$ for addition; skew = .66, $SE = .17$ for subtraction); however, no transformations were undertaken, as this would have hindered interpretation of results. A positive skew in the first year was likely reflective of the fact that the CBM are a) unstandardized measures, and b) many Grade 1 students had not yet acquired

addition or subtraction skills by the fall of the first year. Two outliers among the cases were found for the addition analysis and one was found for both the subtraction and WJ-III analyses, using a $p < .001$ criterion for Mahalanobis distance; because these cases represented less than 1% of the total sample, they were deleted separately for each analysis, leaving $n = 202$ for the addition analysis, and $n = 203$ for the subtraction and WJ-III analyses.

Patterns of missing data were examined using the Missing Values Analysis program in SPSS 19.0.  For the current sample, one child dropped out at the outset of the study (i.e., only demographics were collected), and one child dropped out following the first wave of the study. Seven other children dropped out during study Year-2 (i.e., waves 3 and 4), and one child did not complete the fourth wave of math CBM measures.  Because these missing values represented 5% or less of cases, and because they were characterized as being missing completely at random (Little's MCAR test, $\chi^2 (21) = 18.227$, *ns* for addition, and $\chi^2 (21) = 16.741$, *ns,* for subtraction $\chi^2 (9) = 6.573$, *ns,* for WJ-III fluency), these participants were deleted list-wise from preliminary SPSS analyses (addition CBM, $n = 192$; subtraction CBM, $n = 193$, WJ-III fluency, $n = 194$), although they were retained for HLM analyses as described below.  In terms of parent level of education, 19 cases were missing (approximately 9% of the data). However, most (17 of 19 cases) of the missing parent education data came from a single school, and therefore eliminating these data would have resulted in listwise deletion of a large portion of students from that school. The parent education data was therefore imputed using Hot Deck imputation from the distribution of valid values.  Collinearity tests revealed low levels of multicollinearity among the predictors (all VIF $\leq 2$).

Multilevel analyses were conducted using HLM version 7.0 software (Raudenbush, Bryk, & Congdon, 2010). In the current study, a "cohort-sequential" (Nesselroade & Baltes, 1979) or accelerated longitudinal design was used, in which adjacent segments of limited longitudinal data from different grade cohorts were linked in order to examine a common developmental

trend. In this way, although all participating students had the "potential" to be tested at each of 8 different time points (i.e., twice per year in each of Grades 1 through 4), in reality, each student was only tested at four time points (twice in their respective grades at study entry, and twice in their grade in the second year of the study), as depicted in Table 2.

[INSERT Table 2. *Accelerated growth curve design*]

Each grade cohort contributes a different section to the overall curve while also representing a different pattern of "missingness". HLM was an ideal program for this study design, as it is capable of handling missing data at Level-1 of the model (i.e., the repeated measures within each student). By using the three sets of staggered grade-group data simultaneously, it was possible to build cohort-sequential growth trajectories spanning Grades 1 through 4 (see Figure 1 in Results section).  Given that this approach was effectively linking multiple cohorts, growth within a given cohort may have been due to a developmental process common to each cohort (what is assumed in this design), or to other factors specific to a given cohort. As suggested by Miyazaki and Raudenbush (2000), the appropriateness of an accelerated longitudinal design was determined by comparing the fit of a full model (which included cohort membership as dummy-coded predictors (i.e., each cohort has its own intercept and slope)) to a simpler model where the cohorts are assumed to represent a single developmental trajectory. In order to perform the longitudinal analysis, one copy of the database was restructured from a "wide" data format, where math fluency scores were contained in different variables, to a "long" format, where each math result is a separate case under the same variable (i.e., CBM addition fluency and subtraction fluency, WJ-III fluency). This created the Level-1 database which had an $n = 808$ for CBM addition (i.e., $n = 202$ students at 4 time points), and $n = 812$ for CBM subtraction, and $n = 812$ for WJ-III fluency. The Level- 2 databases were kept in the standard

wide format. Although missing values are permitted at Level-1 of the models while running

HLM analyses, the number of Level-1 cases is limited by the number of valid cases at Level-2.

The 2 students with missing Level-2 data resulted in list-wise deletion, leading to a Level-2 $n =$

200 (total of 2 children missing memory measures as they had dropped out before their

collection) for CBM addition, and $n = 201$ for both CBM subtraction and WJ-III fluency. At

Level-1, there was a total of 22 cases deleted (1 child missing all 4 measures, 1 child who

dropped out after the first round of testing, 7 children who did not complete the two last math

measures, plus one additional case from the student who did not complete the $4^{th}$ round of math

testing), resulting in a Level-1 $n= 785$ for CBM addition, $n = 789$ for CBM subtraction, and $n =$

790 WJ-III fluency.


**Models**

Two-level multilevel models were specified to examine the growth trajectory of math

fluency across Grades 1-4. These same models were subsequently used to examine the effects of

inattention and working memory on the development of addition and subtraction fluency. Level-

1 of the model represents repeated measures within individual students, while Level-2 models

contained the predictor variables for the growth trajectories. In other words, the growth

parameters (i.e., the within-subjects' intercepts and slope) of Level-1 were the outcome variables

to be predicted by the variables at Level-2 (between-subjects).

To determine the best-fitting overall form of the distribution of individual math fluency

growth curves, exploratory models were specified. First, fully unconditional models (i.e., models

with no predictors entered at either Level-1 or Level-2) where no growth was captured (model

assumes variable intercepts, though no change in slope) were examined. This provided a baseline

against which to compare a subsequent model where growth is captured (see Equation 7, below).

Next, a set of polynomial contrasts representing linear, quadratic, and cubic trends were used in

the analysis of each math fluency measure (see Table 3). "Between-grade" contrasts represented

overall changes in skill that occurred from year to year, including the effect of the summer

months (e.g., Grade 1 to Grade 2, etc.). "Within-grade" contrasts depicted growth that occurred

specifically within the (majority of) the school year (i.e., spring to fall, excluding the summer).

Note that the overall between-grade mean was provided by the intercept. Time was modeled in

this way to best capture differences between growth that occurs during the school year as well as

growth (or skill loss) that occurred over the summer months. An example of an unconditional[4]

model with fixed slopes (where each time contrast was significant) would be defined as,


Level-1: (3)

$$Y_{ij} = \beta_{0j} + \beta_{1j}(\text{Between} - \text{Grade Linear}) + \beta_{2j}(\text{Between} - \text{Grade Quadratic})$$
$$+ \beta_{3j}(\text{Between} - \text{Grade Cubic}) + \beta_{4j}(\text{Within} - \text{Grade Mean})$$
$$+ \beta_{5j}(\text{Within} - \text{Grade Linear}) + \beta_{6j}(\text{Within} - \text{Grade Quadratic}) + r_{ij}$$


Level-2: (4)
$$\beta_{0j} = \gamma_{00} + \mu_{0j}$$
$$\beta_{1j} = \gamma_{10}$$
$$\beta_{2j} = \gamma_{20}$$
$$\beta_{3j} = \gamma_{30}$$
$$\beta_{4j} = \gamma_{40}$$
$$\beta_{5j} = \gamma_{50}$$
$$\beta_{6j} = \gamma_{60}$$


where $Y_{ij}$ is the math fluency achievement (CBM or WJ-III) for student $j$ at time $i$. The intercept,

$\beta_{0j}$, (i.e., equivalent to a between-grade mean) was modeled as random at Level-2 ($\mu_{0j}$), to

capture the individual variability in the overall curve level for each student. Parameters $\beta_{1j}$, $\beta_{2j}$,

---

[4] The term *unconditional* refers to models without predictors entered at higher levels. The *fully* unconditional model refers to a model without predictors at either Level-1 (time-variant) or Level-2 (time-invariant). In contrast, interim (Model 0) and final (Model 1) unconditional models refer to models where the developmental trend is specified, but where there are no predictors entered at Level-2

$\beta_{3j}$, $\beta_{4j}$, $\beta_{5j}$, and $\beta_{6j}$ represent between-grade linear, quadratic, and cubic trends and within-grade mean, linear, and quadratic trends, respectively.   Parameter $r_{ij}$ is an error term. Interpretation of the between-grade growth contrasts is similar to what would be considered typical linear, quadratic, and cubic functional forms.  For example, a significant $\beta_{1j}$ parameter would indicate a linear change in math fluency between the grades (e.g., as students move from Grade 1 to Grade 2, etc.). The interpretation of the within-grade contrasts is additive to the between-grade contrasts. Specifically, the within-grade mean was the mean difference between fall and spring testing times within a year (i.e. the net increase in skill that students acquire within the school year), over and above what was predicted from the between-grade growth (i.e., level and slope). It can also be seen as the level of the within-year (spring) curve, representing growth that occurred within a school year that did not occur over the summer. The within-grade linear slope represented any additional linear growth that that occurred when considering the within-grade elements. For example, the within-grade slope could have a steeper pitch than the between-grade one if there was a linear increase in the amount of fluency children learned within a school year with advancing grades that is not accounted for by the between-grade linear slope; this would create a widening effect between within-year (spring) and between-year (fall) slopes, suggesting increasing skill loss over the summer with subsequent grades. Alternatively, a non-significant within-grade linear slope would mean that the within- and between-year slopes are parallel (i.e., end of year curve is simply higher than the beginning of year curve). The within-grade quadratic represented additional instantaneous change in linear slope (i.e., in addition to the between grade quadratic).  See Figure 2 in Results section for an illustration.

[INSERT Table 3. *Contrasts depicting linear, quadratic, and cubic trends, for both between-grade and within-grade growth.*]

These contrasts were entered (uncentered) as time-dependent variables at Level-1 of the model using forward stepping (Nezlek, 2008). Non-significant terms were dropped, so that the interim unconditional (Model 0) and final unconditional (Model 1) developmental models only included contrasts that emerged as significant (although see results on CBM addition). Regarding the error structures, random effects were estimated if they emerged as significant in the preliminary analyses; a more liberal significance level was used ($p < 0.10$), because the coefficients are theoretically random, and to the extent possible, the models should be reflective of this variability (e.g., Nezlek, 2008). Model fit was examined using the log-likelihood ratio test. Therefore, Model 0 represented an interim unconditional model where slopes were fixed, and Model 1 was the final unconditional model where slopes were variable, based on goodness-of-fit tests.

The sign of the correlation between intercept and linear slope variances provided insight into whether fluency development followed a cumulative (positive correlation) or compensatory (negative correlation) pattern (Protopapas et al., 2011). Evidence of a cumulative effect would be seen in the case where variability between children's growth curves increases over time (positive correlation between intercept and slope). Conversely, a compensatory effect would occur in the case of a negative correlation between intercept and slope (i.e., less variability between the slopes over time, as they converge). Correlations between intercept and within-grade mean variances provided insight into whether stronger fluency was associated with greater skill acquisition within a grade.

To address the third research question, working memory and inattention measures were entered independently at Level-2 as predictors of the intercept and each significant growth trend. Sex and parental level of education were entered as control variables at Level-2 of the models. For example, models outlined by Equations 5 and 6 would be specified to test whether verbal working memory, visual-spatial working memory, and inattention predicted both the overall

level of students' growth curves ($\beta_{0j}$) as well as whether these same measures predict linear growth trends ($\beta_{1j}$) while sex and parent level of education were taken into account. In other words, the following equations provide an example of a conditional (i.e., predictors entered at Level-1 and Level-2) model with predictors and varying between-grade linear slopes,

Level-1: (5)

$$Y_{ij} = \beta_{0j} + \beta_{1j}(\text{Between} - \text{Grade Linear}) + \beta_{2j}(\text{Between} - \text{Grade Quadratic})$$
$$+ \beta_{3j}(\text{Between} - \text{Grade Cubic}) + \beta_{4j}(\text{Within} - \text{Grade Mean})$$
$$+ \beta_{5j}(\text{Within} - \text{Grade Linear}) + \beta_{6j}(\text{Within} - \text{Grade Quadratic}) + r_{ij}$$

Level-2: (6)
$$\beta_{0j} = \gamma_{00} + \mu_{01}(\text{Sex}) + \gamma_{02}(\text{Parent Level of Education}) + \gamma_{03}(\text{Verbal } WM)$$
$$+ \gamma_{04}(\text{Visual} - \text{Spatial } WM) + \gamma_{05}(\text{Inattention}) + \mu_{0j}$$

$$\beta_{1j} = \gamma_{10} + \mu_{11}(\text{Sex}) + \gamma_{12}(\text{Parent Level of Education}) + \gamma_{13}(\text{Verbal } WM)$$
$$+ \gamma_{14}(\text{Visual} - \text{Spatial } WM) + \gamma_{15}(\text{Inattention}) + \mu_{1j}$$
$$\beta_{2j} = \gamma_{20}$$
$$\beta_{3j} = \gamma_{30}$$
$$\beta_{4j} = \gamma_{40}$$
$$\beta_{5j} = \gamma_{50}$$
$$\beta_{6j} = \gamma_{60}$$

Level-2 continuous variable predictors were centred around their grand means. Although centering does not affect the estimates of slopes, the interpretation of the fixed intercepts' coefficients is the $\beta$ value when verbal working memory, visual-spatial memory, and attention, were not set to 0, but to their overall mean. Next, forward stepping was used to enter all predictors into the model. Finally, the proportion of the within-student variance that is explained by the fixed effects of time-varying variables (Level-1) was calculated as,

$$\frac{r_{ij}(\text{fully unconditional}) - r_{ij}(\text{Model 1})}{r_{ij}(\text{fully unconditional})} \tag{7}$$

where Model 1 was the final unconditional model (i.e., no predictors at Level-2) in which slopes were treated as random in accordance with goodness-of-fit tests.

The proportion of variance explained at Level-2 (between students) after the addition of predictors at Level-2 was calculated as,

$$\frac{\mu_{0j}(\text{Model 1}) - \mu_{0j}(\text{Model } X)}{\mu_{0j}(\text{Model 1})} \tag{8}$$

where $X$ represented each subsequent model where a new predictor variable was added at Level-2 (i.e., Models 2 through 7 in the Results section, below).

# CHAPTER III: Results

**Growth Curve Developmental Dynamics**

   **Descriptive statistics.** Means and Standard Deviations of CBM and WJ-III fact fluency, for each time point, for each grade-cohort are presented in Tables 4. Correlations between each time point and predictor variables are presented in Table 5 (CBM addition), Table 6 (CBM subtraction), and Table 7 (WJ-III Fluency). It should be noted that in the case of Table 4, the connections of grade, time, and cohort are given explicitly, while in Tables 5-7, the correlations implicitly confound those connections.


      [INSERT Table 4. *Means and Standard Deviations of Math Fluency*]

    [INSERT Table 5. *Correlations Among Predictor Variables and Addition Fluency*]

   [INSERT Table 6. *Correlations Among Predictor Variables and Subtraction Fluency]*

    [INSERT Table 7. *Correlations Among Predictor Variables and WJ-III Fluency]*


   **Accelerated cohort design: Proof of concept.** The appropriateness of an accelerated cohort design was confirmed by testing a full model that included cohort membership as predictors against a simpler model where the cohorts were assumed to represent a single developmental trajectory (Miyazaki & Raudenbush, 2000). Note that this model could only be tested for subtraction; the addition CBM and the WJ-III models included a quadratic, and testing these models led to singularity. The deviance test between the full and simple model was non-significant for subtraction CBM ($\chi^2$ (6, $N$ =789) = 9.15, $p$ = .16). Therefore, the Null Hypothesis that there is no difference between the complex and simple model was unable to be rejected; it was concluded that the simple model was appropriate and that it represented a single developmental trajectory. Figure 1 demonstrates the raw means by cohort and grade corresponding to the accelerated longitudinal design of the current study.

## Specification of Developmental Models

### Determination of the form of developmental trajectories (unconditional models).

The following section outlines the process of determining the interim and final unconditional developmental models for each math fluency outcome (i.e., models with pruned time contrasts and that of random terms according to goodness-of-fit tests). The fully unconditional models (i.e., model with no predictors entered at either Level-1 or Level-2) are reported (FUM) in Tables 8, 9, 10. To specify the Level-1 models for CBM addition, CBM subtraction, and WJ-III, each time-dependent contrast was entered using forward stepping (Nezlek, 2008). For the fixed portion of the analyses, contrasts that did not emerge as significant or did not improve the model fit (estimated by full maximum likelihood) were pruned in order to specify interim unconditional models with fixed slopes (Model 0) and the final unconditional models with random slopes (Model 1), presented in Tables 8, 9, and 10.

Tests of homogeneity of Level-1 variance were non-significant for each model, as required. The reliabilities for these intermediate growth models were .76, .73, .81 for the intercepts and .24, .21, .23 for the slopes (for CBM addition, CBM subtraction, and WJ, respectively); note that relatively low reliabilities do not invalidate the HLM analysis (Raudenbush, Bryk, Cheong, Congdon, & Du Toit, 2004), and the estimates exceeded the minimum threshold of .05 which may indicate that random coefficients should be fixed (Raudenbush & Bryk, 2002). The developmental forms of the trajectories are described according to outcome, below.

### CBM addition. The results contained in the text below represent the statistics obtained during forward stepping of time contrasts while building the developmental models. With

respect to the fixed portion of the addition CBM model, the between-grade linear contrasts emerged as significant ($\beta = 2.31$, $t(584) = 16.60$, $p < .001$), indicating linear growth in addition fluency moving from one grade to the next. The between-grade quadratic term, was also significant ($\beta = -.74$, $t(583) = -3.86$, $p < .001$). The within-grade effects are additive to the between-grade effects; these contrast the growth that occurs specifically within the school years (excluding summer) to the growth that occurs across the grades (including summer). The significant within-grade mean contrast for CBM addition ($\beta = 2.15$, $t(582) = 17.89$, $p < .001$), indicated that there is a mean increase in the level of skill assessed during the school year, parceled out of from the mean level of the developmental curve depicting growth across the years, including the summer months.  The within-grade linear terms were not significant for addition CBM; however, this term was retained in addition model (with no predictors) for accuracy in model specification due to the significant higher-order within-grade quadratic term. The within-grade quadratic term also emerged as significant ($\beta = -.35$, $t(580) = -2.84$, $p < .01$). None of the cubic terms emerged as significant. Each additional time-dependent contrast resulted in a significantly better fit compared to the previous model (each chi-square test was significant at the $p < .001$ level, with the exception of the within-grade quadratic term, which was significant at $p = .03$). This resulted in an interim unconditional model with all significant time-dependent contrasts entered together, though where slopes were treated as fixed (i.e., only the intercept was allowed to vary, as an HLM program default), presented as Model 0 in Table 8.

With respect to random effects, a model where the between-grade linear slopes were allowed to vary between students was a significantly better fit (estimated using restricted maximum likelihood) than one that included only random intercepts for addition ($\chi^2$ (2, $N = 785$) $= 57.77$, $p < .001$). Although random effects for within-grade mean only emerged as significant for the WJ-III (see below), they were modeled as random for each analysis as doing so significantly improved the overall model fit for addition ($\chi^2$ (3, $N = 785$) $= 21.74$, $p < .001$). The

random effect of the within-grade quadratic term was non-significant (each, $p > .50$) and were not modeled in the final analysis; the random effect of the between-grade quadratic term did not reach convergence, and was therefore also excluded from the model. This resulted in a final unconditional developmental model for CBM addition (Model 1 in Table 8).

*CBM subtraction.* Regarding the fixed portion of the CBM subtraction model, the between-grade linear contrast was also significant for subtraction ($\beta = 2.00$, $t(587) = 19.96$, $p < .001$). The between-grade quadratic term was significant when individual slopes were modeled as fixed ($\beta = -.30$, $t(585) = -2.23$, $p = .02$), but non-significant when linear slopes were modeled as random (better fit, as described below); it was therefore dropped from the model.

The within-grade mean contrasts was significant ($\beta = 1.45$, $t(586) = 14.14$, $p < .001$), again, signaling an increase in level of skill assessed during the school year, specifically. The within-grade linear term was not significant and was dropped from the subtraction model. The within-grade quadratic slope was non-significant ($p < .50$) and was also dropped from the model. The cubic term was non-significant. Each additional time-dependent contrast resulted in a significantly better fit compared to the previous model (chi-square test was significant at the $p < .001$ level). An interim unconditional model (with all significant time-dependent contrasts) and fixed slopes (only the intercept varied) is presented as Model 0 in Table 9.

Similar to CBM addition, a model where the between-grade linear slopes were allowed to vary between students resulted in a significantly better fit ($\chi^2 (2, N = 789) = 44.21$, $p < .001$). Again, within-grade mean random effects were modeled as random for the subtraction analysis due to improved model fit, ($\chi^2 (3, N = 789) = 21.39$, $p < .001$). This resulted in a final unconditional developmental model for CBM subtraction (Model 1 in Table 9).

*WJ-III.* With respect to the fixed effects of the WJ-III, the between-grade linear contrast was significant (WJ-III: $\beta = 1.60$, $t(587) = 22.99$, $p < .001$), as was the between-grade quadratic term ($\beta = -.29$, $t(586) = -3.31$, $p < .001$). Similar to the addition and subtraction CBM models,

the within-grade mean contrasts were significant ($\beta$ = .77, $t(585)$ = 11.25, $p$ < .001). In contrast, the within-grade linear, quadratic, and cubic terms were not (p < .50), and were dropped from the models. As with the CBM analyses, an interim unconditional model with significant time-dependent contrasts, and fixed slopes (only the intercept varied) is presented as Model 0 in Table 10.

As in the CBM results, a model where the between-grade linear slopes were allowed to vary between students was a significantly better fit ($\chi^2$ (2, $N$ =790) = 27.29, $p$ < .001). Random effects for within-grade mean emerged as significant for the WJ-III, resulting in a better model fit WJ-III ($\chi^2$ (3, $N$ =790) = 16.81, $p$ = .001).   This resulted in the final unconditional developmental model for the WJ-III (Model 1 in Table 10).

**Final developmental trajectories.** Results of the final multilevel models examining growth trends of fact fluency across Grades 1 to 4 are presented in Model 1 of Tables 8 (CBM addition), 9 (CBM subtraction), and 10 (WJ-III math fluency). Next, a detailed explanation of the developmental trend results is provided, using the CBM subtraction analysis as an example. The subtraction analysis is most straightforward given its linear trend, although a similar rationale applies to both the CBM addition and the WJ-III models; therefore, only noteworthy differences are highlighted here. The significant between-grade linear coefficients indicated linear growth across Grades 1 to 4, meaning that across the years (inclusive of both school and summer months), students' fluency skill increased linearly by an average of 2.04 digits per minute as they moved from one time point to the next across the grades (i.e., 4.08 digits correct per minute, per grade including the summer months).  Further, the significant within-grade mean coefficient indicated that on average, students gained an additional 1.44 digits correct per minute over the school year, which was not gained over the summer months (i.e., slowing of rate of skill acquisition over the summer). In other words, with a linear increase of 2.04 correct digits per

69

time point and an additional 1.44 correct digits correct per minute acquired within a year, the average child would effectively have gained 3.48 digits by the spring testing time (i.e., 2.04 + 1.44); this can be likened to the average linear slope from the fall to the spring (approximately .58 digits correct per minute per month). However, given that children have only gained 4.08 digits correct per minute by the end of summer (i.e., 2.04 + 2.04, and the 1.44 correct digits are now gone), we can calculate the average linear slope from the spring to the fall (summer learning), as 4.08 – 1.44 = 2.64 digits correct per minute (approximately .44 digits correct per minute per month).

In the case of CBM addition and mixed WJ-III fluency, the significant between-grade quadratic trend had a negative valence, indicating a deceleration in the rate at which students acquired fluency over time. Further, for the addition CBM analysis only, the significant within-grade quadratic trend suggests that there was an *additional* quadratic effect (i.e., the within-grade quadratic effect is even more negative than the between-grade effect). In other words, the between-grade trajectory was accelerating faster, but also decelerating faster (i.e., earlier "peak") than the within-grade curve. Note that the between-grade slope is the net effect of the growth that occurs within the school year and the growth that occurs over the summer. Therefore, despite the fact that both the within- and between-year slopes showed a gradual deceleration over time (significant quadratic), children gained progressively more addition fact fluency within each additional grade at a faster rate than they gained over the years, suggesting that students also experienced increasingly slower growth over the summer months with successive grades. Non-significant within-mean linear and quadratic slopes of the CBM subtraction and WJ-III suggest that the ratio of gain within school year to slowing over the summer was constant over the years. The linear and quadratic developmental trends for addition CBM are shown in Figure 2a (Figure 2b shows the trend for subtraction CBM).

[INSERT Figure 2. CBM addition and subtraction fluency growth curves].

Examination of random effects revealed significant inter-individual differences in the level of the growth curve and linear growth rates. In other words, individual differences existed in the overall levels of each math measure fluency curves, as well as in their rates of linear growth. To provide a sense of the distribution of the overall of fluency curves, plausible values ranges for the between-grade trajectories were calculated. For CBM addition, the middle 50% of the children's mean curve levels would fall between 11.12 to 18.98 digits correct (as calculated by Fixed Effect ± SD(.68)). Similarly, 50% of the children's linear slopes would fall between 1.73 to 3.03 digits correct per minute as they moved from one time point to the next across grades. The final developmental trend model explained 57.21% of the within-student (Level-1) variance in CBM addition fluency (calculated using Equation 7, between Model 1 and the fully unconditional model). In the case of CBM subtraction, the middle 50% of the children's mean curve levels would fall between 7.80 and 13.38 digits correct. Similarly, 50% of the children's linear slopes would fall between 1.65 and 2.43 digits per minute as they moved from one time point to the next across grades. Note that the CBM subtraction fluency values are lower than those of CBM addition fluency because, by their very nature, subtraction problems involved fewer correct digits in their solutions than addition problems. The final developmental trend model explained 57.86% of the Level-1 variance in CBM subtraction fluency. Finally, regarding inter-individual variability of WJ-III, the middle 50% of the children's mean curve levels would fall between 5.90 and 12.46 correct responses. The middle 50% of linear slopes would fall between 1.32 to 1.95 digits correct per minute as children moved from one time point to the next across the grades. The final developmental trend model explained 61.38% of the Level-1 variance in WJ-III fluency.

**Longitudinal Growth Patterns (Examination of Cumulative Versus Compensatory Growth)**

Correlations between the intercept variation and slope variation were examined to determine whether growth followed a cumulative, compensatory, or parallel but offset pattern. Positive correlations were found between the intercept variation (i.e., for an average student in Grade 2) and the between-grade linear slope variation for CBM addition (.90), CBM subtraction (.83), and WJ-III (.63). Positive correlations were also found between the intercept variation and the within-grade mean variation for the WJ-III (.40). Overall, these patterns indicate that children with stronger fluency skills experienced faster growth over time compared to those with weaker fluency skills, as well as greater within-year differences; the change in mean was associated with greater individual differences, thereby following a "fan spread" pattern over the grades. Figure 3 (a, b, c) illustrate this point visually. The figures depict math fluency measures (CBM addition, CBM subtraction, and WJ-III) for each of the grades at study entry (i.e., Grades 1, 2, 3). The individual growth curves for each student in the study were drawn using the scale indicated in the lower right corners, and represent the score of each fluency measure according to the four administration times. Each student's growth curve was then plotted according to the mean test score over the four testing points (X-axis) and the between-year growth (i.e., the average of the two upper grade year scores minus the average of the two lower grade scores; Y-axis). Through visual inspection, one can observe that although there is variability between the forms of the curves and rates of growth, overall, for each math measure and each starting grade level, students who have lower mean scores tend to have "flatter" slopes whereas those with higher mean scores tend to have steeper slopes. There also a linear relationship between the level of skill (X-axis) and the between-year growth (Y-axis). Further, looking across the grades (i.e., from cohort 1, to 2, to 3), a clear "fanning out" trend is seen, whereby there is greater variability (curves are more spaced apart) between the levels of the curves as one moves from starting grade to starting grade.

[INSERT Figure 3 (a, b, c). Plot of individual growth curves]

**Individual Differences in Working Memory and Inattention as Predictors of Math Fluency Growth**

To address the third research question, each predictor was entered at Level-2 to determine whether they predicted the developmental trends as established in question 1, and further, whether they explained for variability in the between-student variability in terms of the overall levels and slopes of their growth trajectories.

Tables 8, 9, and 10 display the effects of each predictor on the developmental trajectory of CBM addition, CBM subtraction, and WJ-III fact fluency, respectively.

[INSERT Table 8. *Addition Growth Models with Predictors Entered at Level-2 (Level-1 n = 785, Level-2 n = 200)*]

[INSERT Table 9. *Subtraction Growth Models with Predictors Entered at Level-2 (Level-1 n = 789, Level 2 n= 201)]*

[INSERT able 10. *WJ-III Growth Models with Predictors Entered at Level-2 (Level-1 n = 790, Level-2 n= 201)]*

The final models (Model 6) explained a significant amount of between-student variance in terms of the levels of the overall growth curves (52.85% for CBM addition, 53.32% for CBM subtraction, and 57.45% for the WJ-III). Further, this model accounted for a significant amount of variability between students' rates of linear growth, for both CBM measures (47.36% for addition, and 55.10% for subtraction). In contrast, only 13.04% of linear growth variability in the WJ-III slope was accounted for by the Level-2 predictors.

**Sex**. Sex did not emerge independently as a significant predictor of the overall level of the developmental curve, nor of the linear growth of addition or subtraction. Additional analyses revealed that sex acted as a suppressor variable for inattention, as evidenced by the change in significance of the sex coefficient at the point when this variable was added, by the increased value of β, and by the amount of additional variance explained when sex was included in the model (comparing Models 6 and 7). This is a case of classical suppression, where the effect of inattention on fluency is strengthened, as the variable sex suppresses some of the error variance in inattention. When acting as a suppressor variable, the coefficient for sex is not interpreted (Ludlow and Klein, 2014). Model 7 demonstrated that inattention emerged as significant regardless of the suppression effect. Notably, despite non-significant main effects of sex (addition and WJ-III) or improved model fit, adding it to the model resulted in additional variance explained for both intercept and linear growth for each math measure.

Sex emerged as a significant predictor of the within-grade mean for subtraction (although not addition), such that being male conferred a slight advantage in terms of the average number of digits correct acquired over a school year. Interestingly, however, the fact that sex was not also a significant predictor of the overall growth curve suggests that this advantage is not sustained over time.

**Parent level of education.** Adding parent level of education resulted in improved model fit, although ultimately, it did not emerge as a unique predictor of CBM addition, CBM subtraction, or WJ-III overall level or growth trends once all predictors were entered.

**Verbal working memory**. Results showed a significant positive main effect of verbal working memory on the overall level of each math fluency measure, such that stronger verbal working memory was associated with greater fluency skill (higher trajectory levels). However, verbal working memory was not linked of the rate at which children acquired fluency skills (neither CBM or WJ-III). Adding verbal working memory to the model accounted for 7.30% of

the between-student variance in the CBM addition fluency levels, 7.11% of the variance in CBM subtraction fluency trajectory levels, and 10.75% of the variance in the WJ-III fluency trajectory levels, over and above models that included only control variables.  Verbal working memory also emerged as a predictor of the between-grade quadratic term in the CBM addition analysis. Stronger verbal working memory was related to steeper (more rapid) growth (i.e., higher coefficient resulting in an "earlier peak" quadratic) prior to an eventual deceleration of fluency skill acquisition (as indicated by the negative valence).   Figure 4 depicts the link between verbal working memory on the development of math fluency.

[INSERT Figure 4. Relationship between verbal working memory and a) CBM addition, b) CBM subtraction, and c) WJ-III Math Fluency.]

**Visual-spatial working memory.** There was a positive main effect of visual-spatial working memory on the overall level (intercept) of each growth curve.  Adding visual-spatial working memory to the model accounted for an additional 11.63% in the CBM addition, 12.51% in the CBM subtraction model, and 10.22 % in the WJ-III intercepts over and above models including verbal working memory and the control variables.   Adding visual-spatial working memory to the models also accounted for a significant amount of between-student variability in terms of linear slopes for the CBM models (14.74% for addition, 14.29% for subtraction), although this effect was subsequently accounted for by inattention, as discussed below. Visual-spatial working memory also emerged as the sole predictor of the additional acquisition of fluency skill within each school year, which does not occur over the summer (higher within-grade means) for both CBM analyses. Figure 5 depicts the link between visual-spatial working memory and the development of math fluency.

[INSERT Figure 5. Relationship between visual-spatial working memory and a) CBM addition, b) CBM subtraction, and c) WJ-III Math Fluency.]

**Inattention.** Teacher-rated inattention exerted a significant negative effect on the overall level of the between-grade growth curves. Further, when all variables were entered into the models, inattention emerged as the only significant (negative) predictor of the between-grade linear growth for both CBM operation types. In other words, students who displayed greater levels of inattentive symptoms had math fluency trajectories that were lower in level *and* slower in terms of growth across the grades, compared to peers with who were rated by their classroom teacher as having stronger attentional abilities. Further, as mentioned above, inattention accounted for the variance in linear slopes previously explained by visual-spatial working memory. Notably, adding inattention to the multilevel models accounted for the largest amount of unique variance in each analysis for the overall levels of the CBM addition (27.80%), CBM subtraction (27.66%), and WJ-III math fluency (31.20%). It also accounted for significant unique between-student variance in CBM linear slopes (23.15% for addition; 28.57% for subtraction). Note that although inattention emerged as a predictor of the between-grade quadratic slope, further investigation revealed that inattention was not a predictor of the quadratic function when entered alone, and only emerged as significant when visual working memory was entered into the model. This is another case of a classical suppressor effect, and it is therefore not interpreted or discussed further. Figure 6 depicts the relationship between teacher-rated inattention on the development of math fluency.

[INSERT Figure 6. Relationship between inattention and math fluency for a) CBM addition, b) CBM subtraction, and c) WJ-III Math Fluency.]

**CHAPTER IV: Discussion**

The past two decades have seen increasing interest in the development of mathematical skill. Recent longitudinal studies have provided important insight into the development of general math ability (Aunola et al., 2004; Bodovski & Farkas, 2007; Geary et al., 2012; Morgan et al., 2009; Morgan et al., 2011). By contrast, there is a surprising paucity of longitudinal research focusing specifically on math fluency. The importance of gaining an accurate understanding of this developmental process is highlighted by findings from two previous lines of research: one, which found that math fluency is linked to higher order math abilities (Fuchs et al., 2006; Nelson, Parker, & Zaslofsky, 2016) well beyond the timeframe during which basic arithmetic is emphasized in the curriculum (Nelson et al., 2016); and the other which documented a steady decline in typically developing young adults' ability to solve basic arithmetic problems with both speed and accuracy over the course of recent decades (LeFevre et al., 2014). The results of the study by LeFevre and colleagues are highly concerning, especially in light of the far-reaching health and economic consequences of numerical literacy across the lifespan (Golbeck et al., 2005; Montori & Rothman, 2005; Reyna & Brainerd, 2007; Rivera-Batiz, 1992).

The current study had four major objectives. Using an accelerated longitudinal design, the first aim of this dissertation was to examine the developmental dynamics of basic math fluency (addition and subtraction), simultaneously taking into account the effects of growth within and between the early elementary grades (1-4). This merged and extended developmental research that considers children's mathematical growth trajectories over a number of years, with educational research that examines the slowing/loss of skill acquisition over the summer. The second aim was to determine whether math fluency development followed a cumulative or compensatory growth pattern. The third objective of this study was to investigate the extent to which individual differences in terms of demographics (sex and parental education) as well as behavioral and cognitive constructs (i.e., visual spatial and verbal working memory, and teacher-

rated inattention) predicted developmental patterns and accounted for between-student variability of math fluency skill development. Finally, the fourth objective was to examine differences between addition and subtraction development (in terms of growth curves and predictor variables).

**Developmental Dynamics of Math Fluency**

With respect to the functional form of the growth trajectories, there were both significant between- and within-grade elements; in other words, although there was an overarching shape to the developmental curves across the grades, growth also occurred in a piecewise manner, with greater growth within the school year, and slower growth over the summer months (cf., Cooper et al., 1996; Dawson et al., 2004; Vale et al., 2013). I recognize that the level of the within-grade mean may be somewhat of a conservative estimate, as children in this study were assessed within approximately two months of the end of school and not tested again until approximately two months into the new school year. Given that the beginning of the school year often focuses on review, children may have recouped some of the skills lost over the summer (Allinder & Eicher, 1994). Still, the findings of this study are in line with research that suggests that procedural skills, such as untimed arithmetic, are susceptible to loss in the absence of practice (Cooper et al., 1996). Math fluency may be particularly vulnerable to skill slowing as by its very nature, fast and accurate completion of tasks is dependent on practice (Binder, 1996; Wintre, 1986). Further, to date, most studies on the effects of the summer break, and of summer learning loss that include measures of math achievement have been conducted in the US, although one large-scale Canadian study (Davies & Aurini, 2013), also found evidence of summer learning loss in Ontario schools for both literacy and numeracy. Therefore, the current study also contributes to the literature on the summer setback within the Canadian context.

Differences were found between operation types in terms of the shape of the

developmental trajectories. In the case of CBM addition, children gained skill across the years in a linear manner with a gradual deceleration in the amount of skill children gained with each advancing year (including summer months). When considering the growth that occurs specifically within the school year (i.e., excluding the summer months), the present results also showed that children gained increasingly more skills *within* each school year in the earlier grades, with a gradual deceleration in the amount of skill earned within a year, with advancing grades. Notably, however, the within-year curvilinear trajectory resulted in an earlier peak than the between-year one; by considering the dynamic of the within- and between- year trajectories, it is possible to ascertain a descriptive pattern of the summer effect on skill acquisition. Specifically, in the case of CBM addition, the within-school year curvilinear slope peaked earlier (i.e., showed faster acceleration and deceleration of fluency acquisition) than the between-year trajectory (which represents the net effect of school-year and summer growth), suggesting that children experienced progressively less summer growth with each passing year. Conversely, in the case of subtraction CBM, children gained skills in a linear manner as they moved from one grade to the next, and the rates at which they gained skills within the school year and slowed in skill acquisition over the summer months remained constant from grade to grade (i.e., the within- and between-grade increase in a parallel and linear manner). Regarding the WJ-III (in which there is a mix of addition and subtraction skills (and later, multiplication), between-year growth was curvilinear. However, unlike the CBM addition, the ratio of within-year gains to summer slowing remained constant across the grades. In summary, children's math fluency development progressed in a piece-wise manner, with greater growth during the school years as compared to the summer months. The overarching functional forms across Grades 1 to 4 were decelerating quadratic for the addition CBM and WJ-III measures, and linear for the subtraction CBM measure.

The described differences in developmental trajectories according to operation type may indicate that addition fluency is an easier skill to acquire compared to subtraction (Kamii, Lewis, & Kirkland, 2001) if the assumption is that a decelerating quadratic is an indication of slowing due to gradually closer approximations toward skill mastery. This conclusion is in line with a previous longitudinal study that demonstrated developmental differences between operation types, such that the level of children's addition fluency growth trajectories was consistently higher than that of subtraction fluency, from Grades 1 through 9 (Martens et al., 2011). It may be more difficult for children to acquire fluency in subtraction than in addition if, as previously suggested by Kami and colleagues (2001, 2003), subtraction involves a negative construction of addition, implying that it is contingent on addition skill ability. For example, indirect addition may be used as a strategy for solving subtraction (i.e., subtraction by addition, for example, solving $9 - 7 = 2$, by adding 2 to 7 to make 9), particularly with small differences (De Smedt, Torbeyns, Stassens, Ghesquiere, & Verschaffel, 2010). However, using a nested multiple baseline design, Poncy et al. (2010) failed to find that having fluent addition skills and the conceptual knowledge regarding the relationship between addition and subtraction ("think-addition" strategy) transferred to subtraction fluency skills. This strategy may therefore require explicit instruction of these procedural skills (i.e., direct skill teaching) it does not tend to be learned implicitly through discovery (De Smedt et al., 2010).

An alternative point of view suggests that skill slowing occurs due to fewer opportunities for practice, which could be the case if fluency practice is deemphasized in the curriculum, or if cognitive resources are allocated to more advanced math concepts in the later grades (Binder, 2003). Interestingly, the finding of progressively slower summer growth with advancing years is strikingly similar to the reading results in Cooper and colleagues' (1996) meta-analysis. Although grade was not included as a covariate in the models in this study (as these would have been confounded by cohort), Cooper and colleagues identified a grade effect, such that children

in "fourth grade and beyond showed significant losses (over the summer), some of which were quite dramatic". The authors highlight the counterintuitive nature of the finding, given expectations for factors that would influence children's early school experiences to have the most significant impact on learning. Cooper et al. (1996) proposed a possible floor effect of scaling as an explanation, such that the transformation of raw scores into standard scores eliminates variability present in youngest participants. Using grade-level data, children in grade one can only score one grade below the normed grade level, effectively decreasing the amount of negative change the child can experience. However, as the CBM measures are not norm-referenced, this hypothesis is unlikely to explain the current findings. Although the non-experimental nature of my study does not permit for the determination of causal factors, if fluency skills are not mastered, a shift toward progressively less focus on practicing these skills within the school year may be at the expense of loss of fluency in the absence of practice over the summer. In other words, deceleration may be more reflective of less practice, than of an approach toward mastery.

Of course, if curve form is related to decreased opportunities for practice, it is reasonable to expect the same effect for subtraction. This highlights the further possibility that differences in within- and between- year growth patterns according to operation type reflect distinct problem-solving strategies. Previous research suggests that addition is more likely to be solved through direct retrieval, whereas subtraction is more frequently solved using other algorithmic procedures (Barrouillet et al., 2008). Considering studies that have found accurate retrieval in early grades to predict the use of retrieval in later grades (Bailey et al., 2012), it is conceivable that children could acquire fluency skill based on retrieval (i.e., addition) increasingly more quickly with each passing grade. It is unknown, however, whether problems solved through retrieval (e.g., addition) would be more susceptible to summer slowing in the relative absence of practice, compared to problems solved through an alternative algorithmic strategy (e.g., subtraction). In a

similar vein, differences in terms of susceptibility to summer slowdown have been demonstrated for other math sub-skills, in that calculation is more susceptible to slowing/loss as compared to conceptually based math (e.g., word problems) (Cooper et al., 1996). Therefore, to address these possibilities, future research investigating whether operation types are differentially related to seasonal growth patterns as a function of the strategy used is warranted. Note that the aforementioned possibilities are considered solely at the descriptive level, as it was not possible to perform a direct comparison between operation types. Further, it might be argued that differences in the curve forms reflect a measurement issue (e.g., if probe difficulty was confounded with grade levels). As only one probe per child per time as given, it was not possible to conduct supplementary analyses on whether parallel forms were truly equivalent in item difficulty. However, a decelerating curve was also noted in the mixed WJ-III, where the same test was given at all assessment occasions.

**Longitudinal Growth Patterns (Cumulative Versus Compensatory Growth).**
Regarding the variability of developmental slopes between students, the findings indicated that children differed significantly in terms of the overall level of their growth trajectories, as well as in their rates of growth across the grades. On average, children with higher overall levels of math fluency experienced faster growth over time, as compared to children who started with weaker math fluency skills (cf., Codding et al., 2007; Burns et al., 2010). Additionally, these children acquired greater fluency during the school year (WJ-III). Longitudinally, this resulted in a pattern of increasing differences between individual slopes with advancing grades, creating a fan spread effect consistent with a cumulative model (Matthew effect). These results add to the growing body of longitudinal research that has identified cumulative growth in general math ability (Aunola et al., 2004; Morgan et al., 2009; Morgan et al., 2011) and calculation in terms of accuracy (Salaschek et al., 2014), by demonstrating that similar effects occur in terms of math fluency, both within and across years, and in spite of the summer slowdown. As identified by

Carr and Alexeev (2011), math fluency has a significant impact on the rate of growth of strategy use, from manipulatives (e.g., counting with fingers or counters) to cognitive strategies (e.g., counting mentally or retrieving facts from memory), suggesting that fluency ability itself affects the developmental trajectory. Notably, this effect was strengthened, albeit slightly, by taking into account the predictor variables, and particularly classroom inattention.

**Considerations for developmental research in mathematics: Modeling the summer break.** Specification of accurate longitudinal models is an essential step in understanding developmental processes. Particularly within the context of education, developmental models shed light on various phases of learning (Tenison & Anderson, 2015). Historically, most developmental research has modeled skill growth by assuming smooth trajectories (i.e., most frequently linear and curvilinear). Studies using this modeling approach are unable to account for the effects of instructional breaks, despite the robust finding that the transitional period between grades presents significant deceleration in the learning curve or learning loss, and that this effect is most pronounced in the area of mathematics (Cooper et al., 1996). Therefore, omitting the stable characteristic structure of the school year may paint a potentially incomplete picture of the learning trajectory, if not significantly bias estimates of regression models. In contrast, piecewise models that allow for the estimation of slopes corresponding to different segments of the trajectory are a rarely seen in longitudinal research, although more recent studies have adopted this approach (Kohli, Harring, & Hancock, 2013; Kohli et al., 2015; McCoach & Yu, 2016; Shanley, 2016). Piecewise designs are particularly well-suited to examining portions of the developmental trajectories that deviate from the typical linear or curvilinear forms (i.e., correspond to fundamentally differing patterns of change) (Anderson, 2012; Collins, 2006; McCoach & Yu, 2016), as would be expected when considering growth rates during the school year (instructional period) as compared to the summer months (non-instructional period). In the

current study, the consistent finding of skill slowing over the summer across all three measures of fluency highlights the importance of accounting for seasonal effects in developmental models, as measurement differences are in part dependent on the time of the year children are assessed. Therefore, the results of the current dissertation support the value of models other than linear and curvilinear growth in order to capture developmental change (Kohli et al., 2015), specifically illustrating the utility of a piecewise approach to modeling learning trends that span a number of school years (McCoach & Yu, 2016). While parsimony is desirable, model utility and the ability to inform practice is equally important (Shanley, 2016). The inclusion of seasonal effects in the current study also allowed for the determination of interactions between instructional periods and predictor variables (working memory and inattention), which may have been obscured using smooth trajectory modeling. As discussed below, the differential effect of these predictors on achievement according to time of year may have implications for instruction or interventions. Of course, a major limiting factor to the use of piecewise designs is the sheer number of data points required for accurate modeling (McCoach & Yu, 2016). However, as demonstrated in this study, this obstacle can be overcome using a cohort sequential design (Nesselroade & Baltes, 1979).

**Predictors of Math Fluency Development**

Another important goal of developmental research in the area of mathematics is the identification of individual factors associated with skill growth or deficit, as these may signal flags for weakness in mathematics or opportunities for intervention. When considering the body of research on summer learning, studies have generally been constrained to the realm of educational psychology and policy. The current study contributes to this literature by attempting to merge this area of educational research with cognitive research. The goal was to identify key interactions between seasonal effects (within-year and between-years, including the summer) and individual factors previously found to be associated with math development.

Specifically, although verbal working memory, visual-spatial working memory, and inattention each emerged as independent predictors of the development of math fluency (while controlling for sex and parent level of education), their relative importance differed based on whether the effect was observed within the school year, or across school years including the summer months. These important relationships are masked when within- and between-year effects are not explicitly parceled out. Further, links between seasonal effects and predictors of fluency development highlight implications for the timing of interventions, as discussed further, below. Although the specific cognitive and developmental processes through which these trajectories arise is beyond the scope of the current study, previous research provides a basis for hypothesizing certain possibilities, as discussed next.

**Sex**. Although not the focus of the present study, when examining control variables, there was a slight advantage for boys in terms of acquired subtraction fluency skill within school years specifically. This is in line with previous research demonstrating that boys rely more heavily on direct retrieval strategy than girls when solving basic arithmetic problems (Bailey et al., 2012; Carr & Jessup, 1997). It would follow that more frequent use of retrieval would enhance boys' speeded performance, despite no difference in accuracy (Carr & Jessup, 1997). However, if this were the case, then similar sex differences would be expected in addition fluency. Some studies have suggested that a male advantage in the early elementary grades may be related to visual-spatial ability (Geary, Saults, Liu, & Hoard, 2000; Lummis & Stevenson, 1990). This is an interesting possibility, considering that the association between sex and fluency was found only for subtraction. Previous research has suggested that this operation tends to be solved through quantity-based procedures (Barrouillet et al., 2008; Prado et al., 2014), which have in turn been linked to visual-spatial skills (Geary et al., 2007; Rotzer et al., 2009; Zorzi, Priftis, & Umilta, 2002). An additional analysis that included a sex by visual-spatial working memory interaction effect was non-significant. However, it is noted that in this study, there was not a "pure" measure

of visual-spatial ability, and these skills are separable in terms of their ability to predict number processing proficiency (Krinzinger, Wood, & Willmes, 2012). Therefore, future studies including sex differences and operation types should include such measures. Nonetheless, the within-year sex difference was not maintained across school years (i.e., boys' between-grade slopes were not higher than those of girls, and their between-grade linear growth was comparable), which is in line with other studies that have found no sex differences in children's arithmetic ability (Aunola et al., 2004; Lachance & Mazzocco, 2006; Lindberg et al., 2010). Taken together, the current results suggest that some of the disparity in the literature on the role of sex in math development may be a function of the construct used to measure math ability (e.g., fluency as separate from untimed calculation or word problems, differences in operation type) (Carr, Steiner, Kyser, & Biddlecomb, 2008) as well as the time frame within which math ability is assessed (e.g., no longitudinal effect despite cross-sectional effect). Notably, sex accounted for a significant amount of variance in so far as it was related to other predictors, specifically inattention (Gershon, 2002).

**Working memory.** With respect to the working memory measures, both verbal working memory and visual-spatial working memory were uniquely associated with higher overall slope levels for each fluency measure. This is in line with results of previous studies linking working memory to math fluency (Fuchs et al., 2008; LeFevre et al., 2013; Lukowski et al., 2014; Martin et al., 2014).  Interestingly, differences were found in terms of the relative importance of each working memory measure according to growth phases. Specifically, verbal working memory was predictive of the negative curvilinear trend of CBM addition fluency skill acquisition across the grades, suggesting that children with stronger verbal working had an earlier peak in their curvilinear trajectory compared to children with weaker verbal working memory.  Higher overall curves and earlier peaks in the trajectory may indicate that children with stronger verbal working memory may approach mastery levels more quickly than those with weaker working memory.

87

Alternatively, if the negative quadratic effect were reflective of decreased opportunities for fluency practice with advancing grades, this pattern still suggests that students with stronger verbal working memory acquired skills at a faster rate than their peers with weaker verbal working memory prior to such a shift in instructional focus. Having a solid grasp of quantity and the relationship between numbers is an important precursor to fluency (Baroody et al., 2009). Young children who possess stronger executive functioning skills (including working memory) appear to have an advantage in this area, as recently demonstrated by Fuhs and colleagues (2016). These authors found that Kindergarten students with stronger executive functioning skills had a greater ability to identify number sets (i.e., "sets" of numbers that can be combined to add up to a target value) compared to those with weaker executive functioning skills. In turn, this ability predicted growth in math skill through to the second grade (Fuhs et al., 2016). Working memory may also support fluency development through the maturation of strategies selected to solve arithmetic problems (Geary et al., 2012), from counting to the more sophisticated decomposition strategy.

Visual spatial working memory uniquely predicted the additional skill acquired specifically within the school year, such that stronger visual spatial working memory was associated with higher fluency means in the spring. As first proposed by Heyns (1987), the summer months can be viewed as a temporal control for the effects of schooling. Because students are influenced year round by their non-school environments (e.g., demographics, family, neighborhood, peers, and individual characteristics), taking the effects of the summer into account approximates a natural experiment, thus isolating the effects of formal education when the effects of the non-school environment are held constant (Alexander et al., 2007).

In this way, the link between visual-spatial working memory and math fluency growth occurring specifically within the school year points to the importance of visual-spatial working memory in the process of *active* learning within the school year, and by extension, may be

protective in relation to potential cumulative effects of the summer slowdown. In other words, this finding may represent a longitudinal parallel to previous cross-sectional research that has identified a particularly salient role for visual-spatial working memory during the process of arithmetic skill acquisition (Imbo & Vandierendonck, 2007; Laski et al., 2013; McKenzie et al., 2003; Raghubar et al., 2010). Similar to verbal working memory, visual-spatial working memory may be linked to a more rapid transition to increasingly sophisticated counting strategies (Geary et al., 2007). A specific link between visual-spatial working memory and math fluency may relate to stronger mental models or a better appreciation of the representation of quantity along an internal number line (Berteletti, Man, & Booth, 2015; Geary et al., 2007; Rasmussen & Bisanz, 2005; Rotzer et al., 2009). Visual-spatial skills have been linked to the development of the mental number line (Gunderson et al., 2012; LeFevre, Jiménez Lira et al., 2013), which some authors have found to be linked to calculation (Gunderson et al., 2012), although see LeFevre et al. (2013).  While verbal working memory may take on greater importance as children gain increasing arithmetic proficiency (De Smedt et al., 2009; Holmes & Adams, 2006), visual-spatial working memory skill may continue to be harnessed across the lifespan, depending on the strategy used to solve arithmetic problems (i.e., higher working memory load when using counting compared to memory-based strategies (Hubber, Gilmore, & Cragg, 2014)).

Notably, visual-spatial working memory skill was also associated with of more rapid linear growth in addition and subtraction CBM fluency across years, although this relationship was subsequently accounted for by classroom inattention. Teacher-rated inattention accounting for the relationship between working memory and math ability has been found in previous research (Fuchs et al., 2006). Nonetheless, the finding that visual-spatial working memory remained a significant predictor of within-year growth, despite non-significance over time, is in agreement with results from a short-term longitudinal study from our lab (using a nearly identical sample), which demonstrated that the relationship between boys' inattentive classroom behavior

89

and math fluency (subtraction) one year later was partially mediated by visual-spatial working memory (Gray et al., 2015). Collectively, these findings suggest some overlap between behavioral inattention and visual-spatial working memory. These findings may also be seen as being in line with LeFevre and colleagues' Pathways to Mathematics model, which proposed that spatial attention (indexed by visual-spatial working memory) is linked to children's early math abilities, including numeration, number line understanding, and calculation (LeFevre et al., 2010). Further refinement of this model found that working memory (central executive, visual-spatial sketchpad, and phonological loop) predicted children's arithmetic fact fluency (Sowinski et al., 2015).

Finally, collectively, the differential pattern of verbal and visual-spatial working memory measure according to growth phases supports the view that dynamic shifts in the salience of working memory occur throughout math development, with visual working memory acting as a key predictor of math development in early grades while skills are being acquired, and verbal working memory becoming increasingly important in later grades, perhaps when verbal retrieval may be relied upon more heavily (Holmes & Adams, 2006; Raghubar et al., 2010; Van de Weijer-Bergsma et al., 2015).

**Inattention.** Teacher ratings of children's classroom inattention, at levels well-below those required for a diagnosis of ADHD, were found to be a powerful long-term predictor of mathematical proficiency in this study (cf. Duncan et al., 2007; Garner et al., 2013; Holmberg & Bolte, 2014; Pagani et al., 2010; Pingault et al., 2011; Pingault et al., 2014; Rabiner et al., 2016). Findings add to the burgeoning literature on the deleterious effect of classroom inattention on academic outcomes, by demonstrating that inattentive behavior in the classroom predicts lower overall trajectories and slower growth in math fluency over time for both CBM addition and subtraction. Notably, inattention was the sole unique predictor of growth across grades. In other words, greater levels of inattention are linked to consistently weaker fluency skills over time.

From a longitudinal perspective, the gap in math fluency ability between relatively inattentive children compared to their more attentive peers widens over time. Adding inattention to the model strengthened the relationship between the level of math ability and linear growth, suggesting that classroom inattention may potentiate the cumulative nature of math fluency learning; in this way, children who start out with weaker math skills and who are more inattentive are at a significant risk for chronic weakness in math fluency. What's more, in the current study, teacher-rated inattention accounted for the largest amount of variance, both in terms of the levels of children's growth curves and in rates of growth.

Children with ADHD tend to have chronically low math fluency abilities (Lewandowski et al., 2007; Zentall, 1990; Zentall & Smith, 1993). The current study demonstrated similar findings in a community sample in which most children exhibiting inattention levels well below that of clinical significance (Fuchs et al., 2005). In the larger sample from which our sample was drawn, the proportion of students whose scores on the Strengths and Difficulties Questionnaire (SDQ) were within the at-risk range for inattention/hyperactivity symptoms was approximately 9%, as rated by teachers and parents (Aitken, Martinussen, Wolfe, & Tannock, 2015). Therefore, a sizable minority of students struggle with inattentive behavior that may place them at risk of learning issues. As mentioned above, for the CBM measures, classroom inattention accounted for the same between-student variance in linear growth across the grades that was explained by visual working memory. This finding suggests that children who display greater inattentive symptoms in the classroom may have more difficulty acquiring skills compared to their more attentive peers *regardless* of the time of year. In other words, these children may have difficulty acquiring new skills over the school year, but also gain (or retain) less skill than their more attentive peers over the summer.

The mechanism through which this relationship occurs remains unclear, but here I consider several possibilities. First, behavioral attention is a logical prerequisite to processing

information and benefitting from classroom instruction. The development of attention skills during childhood allows for progressive improvements in children's inhibition skills and their ability to delay gratification (Cerda, Im, & Hughes, 2014; Kochanska, Murray, & Harlan, 2000; Pagani et al., 2010). As proposed by previous authors, an inability to suppress competing behaviors in favor of those that support learning (Duckworth & Seligman, 2006; Pagani et al., 2010) or filter out information unrelated to the instructional task (Li-Grining et al., 2010) might place students at risk for learning difficulties. Given that the time frame of the current study began after the introduction of formal schooling, it is not possible to draw conclusions regarding the directionality of the relationship between inattention and fluency skill development (i.e., whether inattention is a causative agent in fluency underdevelopment, or rather, a symptom of other cognitive weaknesses which in turn predict mathematical difficulties (e.g., number sense, Locuniak & Jordan (2008)).

Nonetheless, a second possibility is that classroom inattention may be reflective of reciprocal/iterative influences. For example, students who demonstrate poorer fluency skills due to, or exacerbated by weaknesses in other cognitive areas that are critical to fluency skill development (e.g., working memory (LeFevre et al., 2013; Martin et al., 2014) or processing speed (Bull & Johnston, 1997)) may also be inattentive. Indeed, in my study, greater inattention was moderately correlated with poorer visual-spatial working memory, and inattention acted as a statistical mediator, where visual-spatial working memory was no longer a significant predictor of the between-grade linear growth (addition and subtraction) once inattention was accounted for. Note that other lines of research have suggested that it is spatial ability specifically (as opposed to purely visual skill, such as recall of static images) that is most closely related to mathematical ability (Passolunghi & Mammarella, 2010). Although my study did not include a measure of static visual memory as a point of comparison, it could be argued that the backward finger windows task involved active manipulation, and therefore required greater attentional

control (Passolunghi & Cornoldi, 2008; Passolunghi & Mammarella, 2010), highlighting a potential overlap between the measures tapping visual-spatial working memory and inattention (Gray et al., 2015; LeFevre et al., 2010). Following a bi-directional model, higher levels of inattention and weaker cognitive ability, could lead to increasingly greater academic difficulty (Diamantopoulou, Rydell, Thorell, & Bohlin, 2007; Thorell, 2007). Indeed, Metcalfe and colleagues (2013) identified a similar reciprocal relationship between inattention and academic achievement (i.e., a composite of reading, writing, and mathematics) in 3- to 6-year-old children.

Third, it is possible that teachers' ratings of inattentive behavior in the classroom may represent a proxy for weak achievement (Fuchs et al., 2006). If this is the case, however, teacher ratings of inattention may still serve as a useful indicator of poor math ability, considering the significant correlations between inattentive behavior and math skill.

Fourth, classroom inattention may be capturing latent constructs, such as academic enablers (e.g., motivation or engagement (Bailey et al., 2014; Plamondon & Martinussen, 2015)). Finally, it is possible that inattention corresponds to a cognitive determinant of math fluency. Although the relationship between behavioral inattention and cognitive attention remains unclear (Gold et al., 2013), mathematical theories suggest a role for attentional resources or attention regulation in fluency development. For example, insufficient allocation of attentional resources to improving the performance of various counting strategies could lead to a delay in memorization of math facts (Shrager & Siegler, 1998), and by extension, to prolonged reliance on inefficient and cognitively taxing counting strategies. Indeed, Geary and colleagues (Geary et al., 2012) found that teacher-rated inattention and intelligence independently predicted the transition to more sophisticated (and therefore more efficient) strategy use in children's addition. Likewise, difficulty inhibiting the use of well-learned but inefficient calculation approaches may impede children's ability to transition to more efficient strategy use, despite having the conceptual knowledge of more effective strategies (Robinson & Dubé, 2013). Further, difficulty

inhibiting irrelevant information from working memory may impede direct fact retrieval (Geary et al., 2012; Passolunghi & Siegel, 2004).

The finding that teacher-rated inattention predicted slower growth over the summer months may be secondary to students' difficulty mastering basic arithmetic within the school year. Fluency theory purports that a benefit of achieving a level of automatic responding is that this ability in itself would be protective in terms of the student forgetting learned facts (Binder, 1996; Mong & Mong, 2010). Students who acquire fluent math facts tend to retain these skills over time, compared to those who learn the skill to achieve accuracy only (Singer-Dudek & Greer, 2005). In other words, skills that are mastered are generalized across time (i.e., maintenance; Poncy et al. (2010)). Although all students may experience a slowdown of skill development in the absence of practice (Cooper et al., 1996), students who have not consolidated math skills during the school year may be at particular risk of skill loss. Alternatively, children who tend to display more inattentive behaviors may have associated cognitive weaknesses (e.g., weak working memory; Gray et al., 2015), which could render them more susceptible to forgetting math facts in the absence of practice.

Of course, these possibilities remain speculative at this time, although they provide a springboard for further study. While issues of etiology remain unclear, the results presented suggest that teachers' ratings of inattentive behavior in the classroom could be used to identify children at risk for long-term weakness in fluency skill, and by extension, general mathematical difficulties (Fuchs et al., 2006; Hecht et al., 2001). Consequently, these findings also point to the potential use of accommodations or early interventions for children displaying classroom inattention, as discussed further below.

Seems like you need a new heading here (next part is not about inattention but seems to be about limitations of the measures).

With respect to the WJ-III, no predictors accounted for the variance in slopes. One possible explanation for this finding relates to the measures themselves. Although the WJ-III and M-CBM both contain only basic facts, on the WJ-III, problem stems range from 0 to 10 (rather than 0 to 12 in M-CBM), and problems are arranged in order of increasing difficulty (in terms of operation type (i.e., addition and subtraction, and then multiplication)) and response requirements (i.e., the first 31 problems are single-digit responses, followed by a mix of single and two-digit responses). Therefore, it may be that more children were able to respond to the earlier questions fairly consistently on the WJ-III. Conversely, on the M-CBM, encountering relatively more challenging questions earlier on in the probe (e.g., one requiring a two-digit response), may differentiate weaker from stronger children (with the latter being able to obtain more digits correct early on, in this example), thus leading to greater variability (range) in responses for the sample. Inspection of the variance components of the unconditional models indicated less variation on the WJ-III ($SD = .47$) compared to CBM addition ($SD = .97$), and CBM subtraction ($SD = .70$). Therefore, changes in variance may have been more difficult to detect on the WJ-III due to the limited variability of slopes in the first place (e.g., Nezlek, Kafetsios, & Smith, 2008).

Alternatively, it could be argued that this difference across tasks may indicate that M-CBM fluency growth is reflective of measurement error, rather than actual skill level, which limits the use of M-CBM probes for the tracking of skill growth. As has been noted by previous authors, repeated administrations of CBM can produce variable estimates of response rates depending on the type of measure used (Christ & Vining, 2006; Hintze et al., 2002; Methe et al., 2015). Thus, although the CBM captures gradual skill improvement (Prindle et al., 2016), a potential trade-off is a reduction in psychometric reliability and score equivalence, because alternate use of forms, multiple testing occasions and different raters introduce potential sources of error (Christ, Van Norman, & Nelson, 2016). However, this issue appears to be more relevant

to multi-skill math CBM probes (Christ & Vining, 2006; Methe et al., 2015), whereas 2-minute administration of probes constrained to a single skill (such as those used in the current study) have been shown to provide reliable, valid, and dependable assessment of ability using a single probe (Hintze et al., 2002). Further, the fact that similar predictor patterns were seen in both addition and subtraction analyses in terms of slope, and between all three measures in terms of overall curve level, suggests that the results of the current study are quite robust.

A final consideration is that, despite the important contribution of basic math fluency in the development of higher-level math skill, arithmetic fluency is a narrow skill that is only moderately correlated with other math abilities; in other words, it is not in and of itself representative of more advanced mathematical skill (Nelson et al., 2016). As such, longitudinal studies that investigate the development of fluent performance in more complex math skills (e.g., multi-step algorithmic computation) are indicated. An important next step, therefore, is further psychometrically-based research that ensures adequately equated forms, both horizontally (within a grade) and vertically (across grades), so that these measures can be used with greater confidence to assess fluency development over time.

**Educational and Clinical Implications**

Results of the current study highlight potentially important implications for educational practice. First, as mentioned above, the possibility that the developmental trajectory of subtraction fluency lags behind that of addition would indicate that a focus on a mastering addition may also support the acquisition of subtraction fluency, as previously suggested by Kamii and colleagues (Kamii & Lewis, 2003; Kamii et al., 2001), although this may require direct skill instruction (i.e., rather than relying on self-directed discovery; De Smedt et al., 2010). Second, children who experience significant weakness in math fluency improve less than more fluent children over time, perhaps in part because children respond differentially to instructional

approaches depending on their level of proficiency with number combinations (Codding et al., 2007). Importantly, the current research suggests that chronic difficulties in basic math fluency appear to be compounded by inattentive behaviors in the classroom. With this pattern in mind, behavioral inattention, even at levels well below those required for the diagnosis of ADHD, predicts a profound and lasting negative impact on a child's developmental trajectory in math. What is more, in my study, the seasonal effect was differentially related to cognitive and behavioral domains, such that the slower growth in math ability seen in relatively inattentive children extends to the summer months.  As such, the results presented provide the basis for hypothesizing that ratings of behavior may be useful in screening for mathematical difficulties and may identify children who may benefit from early intervention, whether within the school year or over the summer holiday. It is also possible that inattentive behaviors themselves may present a target for academic remediation. Alternatively, latent constructs (e.g., motivation) that may be captured by teacher-rated inattention may be important to target (Bailey et al., 2014). Finally, stronger visual working memory predicts greater skill acquisition within the school year, which may also be protective in terms of the effects of the summer (e.g., recoupment of skill). Further research is needed to explore these possibilities. In the remainder of the chapter, I review issues related to fluency instruction and remediation, and discuss how these may relate to my findings.

**General Implications for Educators and School Psychologists.** The cumulative nature of math fluency development highlighted in the current study suggests important practical implications for educators and school psychologists. Historically, in the Ontario education system (where this study took place), children referred for a psychological assessment have already fallen well behind curricular expectations (often within the range of a one to two-year lag), which can subsequently lead to formal access to special education services. Further, students who may benefit from fluency instruction or intervention are unlikely to be limited to

those who struggle to such a degree where a comprehensive psychological assessment is warranted. Teachers are therefore well-positioned to identify children at risk for academic difficulties and implement effective intervention strategies. As highlighted by Methe and colleagues (2012), timely classroom-based intervention may offset the need for more intensive services (Ysseldyke, Vanderwood, & Shriner, 1997) or mitigate the development of more significant math difficulties (Mazzocco, 2007). Further, the benefits of fluency intervention have been shown to generalize to higher-level math skills such as computation (McTiernan, Holloway, Healy, & Hogan, 2016). Therefore, focusing on instructional strategies that are empirically sound, cost-effective, engaging, generalizable, that can be easily implemented, and that may benefit all children but that can also target specific weakness in fluency, are of the utmost importance.

**Considerations regarding instructional and intervention strategies**. Despite the far-reaching influence of math fluency skill across the lifespan, there is a surprising lack of focus devoted specifically to math fluency instruction and remediation in the literature (Codding, Hilt-Panahon, Panahon, & Benson, 2009; Codding, Burns, & Lukito, 2011). Many math intervention packages focus on improving mathematical knowledge (e.g., number sense, algorithms; Clarke et al., 2016). Further, early meta-analyses evaluating the effectiveness of math interventions have not considered fluency as separate from other math skills (e.g., Swanson & Sachse-Lee, 2000). Of the studies that have focused on math fluency interventions specifically, the majority have utilized small n, single-subject designs (Reisener et al., 2016; Whitney, Hirn, & Lingo, 2016), and these have mainly targeted the acquisition/frustration phase of learning (i.e., when children are first learning to complete a skill accurately), as opposed to learning phases where increasing fluency or skill generalization are primary goals (Burns, Codding, Boice, & Lukito, 2010; Poncy et al., 2010). Indeed, in a meta-analysis summarizing interventions supporting students struggling with mathematics, Codding and colleagues (2009) found that 68% of studies

employed a single-subject design. Similarly, there are no randomized control trials (RCT) specific to fluency intervention in the literature (Codding et al., 2007; McTiernan et al., 2016). While single-subject designs are important as they can be highly effective at determining treatment efficacy at the individual level, they should not be used in isolation to make determinations related to policy or practice, due to their lack of generalizability (Kavale & Forness, 2000). To this end, a few studies have taken a meta-analytic approach to analyze single-subject studies (Burns et al., 2010; Codding et al., 2011; Joseph et al., 2012), in order to help establish external validity. The discussion below focuses on three main issues emerging from extant literature which may inform practice: key components of instruction/intervention, interactions between students' skill level and intervention strategy, and the timing of interventions. Further considerations regarding the relations among inattentive behavior, fluency interventions, and instructional timing are also discussed.

 **Considerations regarding intervention components.** Children at risk for weaknesses in mathematics may benefit from direct training in fluency (Hartnedy, Mozzoni, & Fahoum, 2005; McTiernan et al., 2016). Components of fluency training may include (a) explicit instruction regarding the target skill using overt responding techniques (e.g., see/say), (b) the use of manipulatives, (c) drill (learning facts in isolation (Haring & Eaton, 1978)), (e) practice (the use of learned responses in combination with previously learned responses), (f) corrective feedback, review, written or computerized exercises, (g) self-management and (h) motivators to maintain engagement to with the task (Burns et al., 2015; Clarke et al., 2016; Clarke, Doabler, Nelson, & Shanley, 2015; Fuchs, Powell et al., 2010). A central component of fluency skill training involves practice, typically, repeated learning trials to develop speed and accuracy (see Clarke et al., 2016 for a review). However, not all practice opportunities are equal. Codding and colleagues (2011) synthesized 17 single-subject design studies to analyze key treatment components, treatment intensity, and feasibly of treatment delivery. They found that

interventions that combined multiple components (three or more), such as drill, modeling with practice, and self-management strategies, were more effective than those with fewer elements. Further, a critical finding was that drill and practice that also includes demonstration (i.e., modeling, where students observe the process or steps of the skill to be learned; this may be self- or teacher-directed) emerged as an effective remedial strategy (mean *Phi* = .71) for children with fluency weaknesses, although drill and practice alone was not (mean *Phi* = -.003), in support of the view that practice is a necessary but not sufficient aspect of math fact instruction (Baroody et al., 2009; Binder, 1996; Daly, Martens, Barnett, Witt, & Olson, 2007; Rivera & Bryant, 1992).

Children need a fundamental understanding of number combinations (Burns et al., 2015), because weaknesses in conceptual understanding of basic mathematical concepts (e.g., inversion, commutative properties) may impair procedural fluency (i.e., knowledge of rules, symbols, and steps required to solve math problems; Burns et al., 2015). Therefore, a focus on fluency should not come at the expense of conceptual understanding (or vice versa) (Villasenor & Kepner, 1993). Rather than being mutually exclusive, fluency and conceptual understanding emerge from an iterative developmental process (Clarke et al., 2016; Rittle-Johnson, Siegler, & Alibali, 2001; Rittle-Johnson, Schneider, & Star, 2015). For example, in a RCT with 178 Grade 2 students, Carr et al. (2011) found that students who received both conceptual instruction as well as fluency training performed significantly better than those who were in a control group (no math intervention), and those who received other fluency or conceptual training only. Indeed, previous research on curriculum development suggests that an emphasis on mastery of basic skills within a rich problem-solving based context presents an effective approach to developing computation abilities (Reys, Reys, & Koyama, 1996).

***Examples of intervention approaches.*** One example of a repeated-trial fluency training intervention which includes drill and practice with modeling, Cover-Copy-Compare (CCC) (Skinner, Turco, Beatty, & Rasavage, 1989), is prominently featured in the fluency intervention

literature (e.g., 7 out of 17 studies including in the Codding et al., 2011 analysis were CCC or a variant). In CCC, fluency develops through a combination of repeated exposures to problems and their answers, repeated opportunities to respond, and immediate feedback to ensure accuracy and avoid the reinforcement of errors (Carr et al., 2011). Students move through five steps, including viewing the math fact and solution (written on left side of the page), covering the math fact and solution, writing the fact and solution (on the right side of the page), uncovering the original fact/answer combination, and comparing their response to the model. Literature reviews and meta-analyses of single-study research have identified CCC and its variants as an effective strategy for building math fluency in typically-achieving children, as well as in students with difficulties and disabilities in mathematics (Codding et al., 2011; Joseph et al., 2012; Stocker & Kubina, 2016).

CCC is simple and cost effective (i.e., is a non-commercial product). Students can easily learn this strategy at school, which can then be reinforced by parents (Stocker & Kubina, 2016). On the other hand, drill and practice with modeling strategies have been criticized for not engaging students, and treatment effectiveness is limited by whether students will adhere to the program (Hawkins et al., 2016). Experimental studies have found that technology-delivered fluency intervention, or computer-assisted instruction (CAI; through computers, as well as mobile applications "apps"), may provide an effective and cost-effective alternative (Burns, Kanive, & DeGrande, 2012; Ysseldyke, Thill, Pohl, & Bolt, 2005). These programs, which incorporate drill and practice, as well as self-monitoring, have become increasingly popular (Hawkins, Collins, Hernan, & Flowers, 2016). Recently, Hawkins and colleagues (2016) provided guidelines to help teachers implement computer-assisted instruction. These guidelines include similar components to those in traditional methods, such as opportunities to respond (i.e., repeated practice), immediate feedback (with the option for overcorrection, or the opportunity to

repeat correct response following an incorrect one), and pacing that is fast enough to maintain student engagement but appropriate to the instructional level.

Results from a meta-analysis of computer-based interventions (Cheung & Slavin, 2013), indicated that CAI are an appropriate adjunct to the math curriculum (ES = +0.15 for CAI and ES = +0.19 for CAI when used to supplement math instruction). The effectiveness of this approach is highlighted by an experimental study that included children in Grade 3 and 4 who were initially categorized as being at risk of serious math difficulties or disabilities (i.e., scoring below the 25th percentile on a group administered math achievement test) (Burns et al., 2012). The authors found that significantly fewer students who had received the computer-based fluency intervention were still categorized as at-risk for math failure at the post-test as compared to the control group. Not all students respond favorably to computer-based programs, and other work suggests that some students respond better to more traditional methods (Cates, 2005; Mautone, DuPaul, & Jitendra, 2005). Further, Cheung and Slavin (2013) found that CAI does not provide a replacement for sound teaching practices (ES = + 0.06 when used as a core component to math instruction). One limitation of CAI methods is that they focus only on developing fluency, which may not be appropriate for students whose struggles also reflect difficulties in conceptual understanding (Burns et al., 2012).

My study highlighted visual-spatial working memory as a correlate in the development of math fluency, which raises the issue of whether working memory training might benefit students' acquisition of fluency. The basic premise of working memory interventions is that targeting domain-general factors variables that are correlated with learning will translate to academic skill growth ("far transfer"). However, meta-analyses (Melby-Lervag & Hulme, 2013; Melby-Lervag & Hulme, 2016; Melby-Lervag, Redick, & Hulme, 2016) have failed to find convincing evidence supporting the improvement of academic skills after working memory training (Melby-Lervag & Hulme, 2013; Melby-Lervag et al., 2016; Redick, Shipstead, Wiemers, Melby-Lervåg,

102

& Hulme, 2015; Titz & Karbach, 2014). For example, in a recent randomized control trial, Roberts and colleagues (2016) found that although working memory training in 6- and 7-year-old children resulted in short-term improvements in visual-spatial working memory at one year, these effects were not sustained at two years. Further, these authors found no improvement in math calculation, and in fact, observed a negative association between working memory training and math outcomes two years post intervention.

Working memory training programs may be more effective for younger children. For example, Passolunghi and colleagues (2016) found a link between working memory training and early math abilities/conceptual understanding of numerical quantities in preschool children. Still, other studies that have considered young children found that working memory training was not as effective as training in counting for increasing early numeracy skills (Kyttala, Kanerva, & Kroesbergen, 2015). Therefore, considering the significant cost of working memory training programs (Melby-Lervag & Hulme, 2013; Melby-Lervag et al., 2016; Redick et al., 2015; Titz & Karbach, 2014), and the lack of evidence for transfer to specific math skills, these programs would at a minimum require further study prior to implementation.

Other traditional strategies with demonstrated effectiveness (see Clarke et al. (2016), for a review) include Incremental Rehearsal (Burns, 2005), which is a drill flashcard based intervention without conceptual teaching, and Taped-Problems (McCallum, Skinner, & Hutchins, 2004; Poncy et al., 2012), where the student attempts to answer problems before the answer is provided on an audiotape. Multi-component interventions include Detect, Practice, and Repair (Poncy, Skinner, & O'Mara, 2006), Math to Mastery (Doggett, Henington, & Johnson-Gros, 2006), and Great-Leaps Math (Mercer, Mercer, & Campbell, 2002; Whitney et al., 2016). Next steps will be to compare treatment strategies in classroom settings using RCT methods. For example, using a classroom-wide approach, Poncy and colleagues (2012) found that the Taped Problems strategy led to greater increases in arithmetic skills compared to CCC

for many grade 3 children, when presented as class-wide interventions. However, Poncy et al. found that a minority of children responded better to CCC, suggesting differences in student responses to intervention. In the next section, I discuss one of the factors identified in the literature to consider when adapting effective interventions, the skill-by-intervention interaction.

**Considerations regarding skill-by-intervention interactions (learner characteristics).** When determining the utility of fluency instructional approaches or interventions, an important factor to consider relates to matching students' relative skill level to an intervention strategy that meets the need of that student. Research in this area has used the *instructional hierarchy* (Haring, Lovitt, Eaton, & Hansen, 1978), which describes students as moving through four stages of increasing skill (i.e., acquisition, fluency, generalization, adaptation), progressing from slow and deliberate responding, to fluent with the ability to utilize learned skills in novel settings. A similar hierarchical concept is that of *instructional levels* (Burns et al., 2006; Gickling & Thompson, 1985), also referred to as benchmarks in CBM (Hosp et al., 2007). Here, observable behavior (e.g., digits correct per minute) is used to categorize students by skill level ranges.

Studies have shown that an appropriate match between student skill level and intervention strategy optimizes treatment effects (Burns et al., 2010). Specifically, fluency skill that falls below 14 digits correct per minute (DCPM) for students in Grades 2-3, or below 24 DCPM for 4th and 5th grade students, is referred to as the "frustration" range of performance (Burns et al., 2006) (i.e., similar to the "acquisition" stage of the instructional hierarchy). At this level, the primary goal is for the development of accurate responding, and the intervention need is scaffolding and direct instruction. Students are considered to be within the "instructional" range of performance (i.e., similar to "fluency" stage of the instructional hierarchy) when fluency levels fall between 14-31 DCPM in Grades 2-3 or 24-49 DCPM in Grades 4 and 5; for these children, the goal is increasing fluency and the suggested interventions include drill tasks (Burns et al., 2006). Finally, the "mastery" range is reached when children's fluency is > 31 DCPM

(Grade 2-3) or > 49 DCPM (Grades 4-5). Here, the learning goal is for generalization and the suggested approach is to encourage problem solving with diverse and novel items and applications (corresponding to "generalization" and "adaptation" stages of the instructional hierarchy; Burns et al. (2006)).

According to the instructional levels outlined by Burns et al. (2006), the mean DCPM in my study (addition), suggests that most students fell within the frustration range in the earlier years, with progression toward the acquisition range. However, the variability of the results suggests that at any given grade level, some students are particularly strong (i.e., at or close to the next instructional level), whereas other may lag behind (e.g., still at the frustration phase when same aged peers are on average working within the acquisition range). Therefore, these results point to the need for a class-wide instructional approach which is flexible enough to also meet the needs of individual students.

In one of the few randomized control trials investigating the effectiveness of fluency interventions, Codding et al. (2007) found a significant interaction between initial level of fluency skill and response to intervention, such that students who experienced fluency levels within the *frustration* range benefitted most from an approach that included modeling (Cover-Copy-Compare; Skinner et al., (1989; 1997)), whereas those students whose fluency levels fell within the *instructional* range benefitted most from an intervention focused on drill within a time limit (Explicit Timing or ET; Rhymer et al., 2002). The skill level by treatment interaction was further examined using meta-analysis by Burns et al. (2010), who conducted a meta-analysis of single-study research which included instructional approaches/interventions other than CCC and ET. These authors found that in Grade 2 to 6 children, interventions targeting initial acquisition of skills resulted in large effect sizes for those at a frustration level, but only moderate improvement for students at an instructional level. Although the authors were unable to draw firm conclusions regarding the effect of fluency interventions (i.e., to be used for those with

higher proficiency because of the small number of studies that included students with more advanced skills), they noted that fluency interventions had only small to moderate effects on students with weaker skills (i.e., those at the frustration level).

Similar results have been demonstrated using single-study designs in recent years. Using a single-case design (multiple baseline with two first-grade students and one third grade student), students who struggled with conceptual knowledge received an intervention to address this weakness (i.e., modeling), whereas students who demonstrated weakness in fluency (but who showed acceptable knowledge) were administered an intervention focusing on procedural fluency without conceptual teaching (incremental rehearsal) (Burns et al., 2015). Following the establishment of a baseline, the contra-indicated instruction procedure was administered for the first two weeks (i.e., those with conceptual deficits received fluency training, and vice versa); the prescribed intervention was administered for the following two weeks. Results indicated that the prescribed (well-matched) intervention was significantly more effective than the contra-indicated intervention regardless of which intervention this was. The percentage of non-overlapping data (PND; non-parametric statistic providing a measure of how many points of the intervention phase fall above the highest point of the previous phases) ranged from 70 to 100, indicating that 70-100 percent of the students' data during the experimental phase fell above the highest point of the previous phases. These findings are in line with an earlier non-experimental study (i.e., no third staggered implementation phase recommended to demonstrate experimental control; Burns, 2011).

Taken together, these studies demonstrate the utility of ensuring that the intervention is well matched to the students' level of performance. Indeed, these studies provide an illustration of the notion of a "window of learning" (Tucker, 1985), where learning is optimized when it occurs at a level that is challenging enough to stave off boredom, but not so difficult as to be frustrating (Burns et al., 2006).

106

**Considerations regarding the timing of instruction/interventions.** In addition to treatment effectiveness (i.e., whether an intervention is effective at increasing learners' skill levels), an important practical consideration for educators is treatment efficiency (i.e., how well or how quickly it works). In a recent meta-analysis, Poncy and colleagues (2015) highlight the issue of an intervention's *relative* effectiveness, by noting that remedial approaches may be similarly effective (change the level of ability), although they may not be equally efficient (taking into account the rate of improvement). By accounting for both the time devoted to the intervention and the scope of the intervention material presented, these authors established that explicit timing was twice as efficient as incremental rehearsal in improving fluency skill when basing the analysis on learning rate, despite the finding that effect sizes of learning (net) suggested that incremental rehearsal was more effective than explicit timing.

Another timing consideration with respect to fluency building relates to the organization of practice when the total time is held constant. Schutte and colleagues (2015) found that third grade students made greater gains in addition fluency with distributed practice (i.e., 1-minute practice 4 times per day) than with massed practice (i.e., once per day for 4 minutes). This finding is in line with the well-established "spacing effect" seen in memory and learning research, where retention is improved when presentation of learned material is spread out over time (see Cepeda, Pashler, Vul, Wixted, & Rohrer, 2006 for a review). These studies have implications for teachers, as they suggest that optimized outcomes do not have to come at the expense of additional time or resources. Note, however, that few studies have studied timing and relative efficiency (i.e., comparing the effectiveness of interventions), and thus further research in this area is needed.

Results of my dissertation also suggest that the timing of potential interventions is an important consideration. First, consistent with previous longitudinal studies of mathematics in general and of fluency skills specifically (cumulative results in this study), early interventions are

107

indicated.  However, research regarding the timing and frequency of interventions is needed, as previous studies have noted that the effect of early interventions often "fade out" over time (Bailey et al., 2014; Cooper, Charlton, Valentine, & Muhlenbruck, 2000; Cooper, Valentine, Charlton, & Melson, 2003). This pattern suggests that one-time early intervention may be insufficient to support children at risk of math weaknesses, and that close monitoring, repeated interventions, or implementations of supports targeting other stable latent traits that also affect math ability such as motivation (e.g., Bailey et al., 2014; Plamondon & Martinussen, 2015) may be indicated. Considering the results of the current study, one approach would be the provision of intervention over the summer months.

Current initiatives in Ontario are focused on providing short-term (3-week) intensive interventions over the summer months to elementary students (Grades 1-5) who were identified as possibility being able to benefit from instruction to offset learning loss (Davies & Aurini, 2013). This initiative started in 2010 and was coordinated by the Council of Ontario Directors of Education (CODE). It was funded by the Literacy and Numeracy Secretariat, Ontario Ministry of Education, and was initially directed at reading interventions (Davies & Aurini, 2010). Since 2013, the province has expanded this approach to address weakness in mathematics. As this is a relatively recent, ongoing large-scale study, published data is limited, although initial results suggest promising outcomes (Davies & Aurini, 2014; Davies, Aurini, Milne, & Jean-Pierre, 2015).  Specifically, Davies and Aurini (2014) reported that although both attendees of the numeracy program ($n = 463$) and controls ($n = 1,888$) experienced summer learning losses, attendees of the program fared better, showing one month more skill compared to the control group (reported effect size = .17), when student characteristics (prior numeracy, grades, and SES factors) were taken into account. In essence, this short-term intervention effectively offsets some of the summer losses, thus allowing children to "hit the ground running" when school resumes in

the fall (Davies & Aurini, 2014). Qualitatively, the authors also describe increased engagement with the material. However, interactions with symptoms of classroom inattention are unknown.

This line of research and intervention is important, not only in because it can offset summer losses, but also because it capitalizes on teaching resources by reducing the amount of time needed for review and re-teaching the previous year's content. A national U.S. survey of 500 teachers suggested that a majority (66%) spend at least 3-4 weeks re-teaching material from the previous year (National Summer Learning Association, 2013). Further, in terms of review strategies specific to arithmetic, in light of the summer slowdown, it is conceivable that some students will begin the year falling within a different instructional level from that which he/she was at in the spring. For example, consider a child who is performing at the instructional level at the end of Grade 3, but who is functioning at a frustration level for Grade 4 skills. It is unclear whether this student would benefit from further instruction with modeling, or whether they would simply require drill and practice. Studies indicate that skills are rapidly recouped after the summer months (Allinder & Eicher, 1994), in support of a focus on drill and practice, however, some combination of instructional methods might constitute the most effective review for many children.

**Consideration of the relationship between inattentive behavior, fluency interventions, and instructional timing.** Although questionnaires related to symptoms of inattention (and hyperactivity) are most often administered within the context of an assessment regarding a potential diagnosis of Attention-Deficit/Hyperactivity Disorder, the current results point to the use of these tools as a potential screener for difficulties in basic academic skills (in the present case, math fluency). For example, irrespective of diagnosis, high levels of inattentive behavior may identify children who may be at risk for early academic difficulties. This identification is particularly relevant for children who display a greater number of inattentive behaviors, compared to children who also display hyperactive/impulsive behaviors, because

109

inattentive children may be less likely to come to the attention of school professionals (Willcutt, 2012).

Research considering accommodations for behavioral concerns has been conducted using clinical samples (i.e., diagnosis of ADHD), but as outlined by Harrison and colleagues (2013), evidence for the utility of classroom accommodations for ADHD (e.g., additional time, preferential classroom seating, etc.) on academic outcomes is surprisingly lacking. A commonly suggested accommodation for children struggling with dysfluent responding is the allocation additional time to complete tasks. Lewandowski and colleagues (2007) found that extended time limits on math fluency tasks may allow students diagnosed with ADHD achieve to their potential without conferring a distinct academic advantage over their typically-achieving peers, thereby possibly mitigating the impact of disability associated with ADHD.  Other studies, however, have not found that additional time enhances performance in children with attentional difficulties (Pariseau, Fabiano, Massetti, Hart, & Pelham, 2010), and highlight the counterintuitive nature of this recommendation given that children with ADHD often struggle to sustain attention during prolonged periods of time (e.g., Barkley, 1997). Therefore, it is currently unknown whether classroom accommodations could appropriately support students displaying classroom inattention at risk of math weakness.

Results of my study indicate a close longitudinal link between classroom inattention and fluency development, suggesting the importance of considering both factors when selecting fluency interventions. Specifically, one might postulate that if interventions delivered during the school year are effective at increasing fluency growth, these may be protective in terms of skill slowing or further loss that inattentive children may face over the summer.

Targeted training in fluency may be beneficial for children who struggle with fluency and who concurrently display symptoms of inattention. In a general sense, research suggests that the ability to respond fluently during tasks has the associated benefit of building children's task

endurance (i.e., maintenance of a certain level of performance, or attention span, on a given task) (Binder, Haughton, & Van Eyk, 1990). Other lines of evidence suggest that self-management of practice is associated with increased attention, motivation, and independence (Reid, Trout, & Schartz, 2005). Results from a meta-analysis (Codding et al., 2011) demonstrated that student-managed fluency interventions (i.e., interventions that are self-directed) resulted in large effect sizes for treatment effectiveness, which as noted by the authors, highlights the beneficial nature of students taking responsibility for their own learning (McDougall & Brady, 1998). Further, research on technology-based interventions also suggests concurrent improvements in both math fluency and on-task behavior (Burns et al., 2012; Mautone et al., 2005). In other words, following a model of reciprocal influences described above, it would be conceivable that fluency developed through targeted training programs could engender improvement in classroom inattentive behavior.

Students whose mathematical development is impeded by inattentive behavior may also benefit from interventions where behavioral symptomatology and academic functioning are concurrent targets. For example, class-wide peer tutoring (CWPT) which integrates one-to-one peer tutoring, frequent immediate feedback (Duhon, House, Hastings, Poncy, & Solomon, 2015), active student responding, and positive reinforcement, has been found to effectively increase both on-task behavior and academic productivity in children with ADHD (Buzhardt, Greenwood, Abbott, & Tapia, 2007; DuPaul & Eckert, 1998; DuPaul, Ervin, Hook, & McGoey, 1998; Maheady & Gard, 2010; Raggi & Chronis, 2006). Importantly, CWPT utilizes a game format, whose competitive nature may build motivation for math tasks (Plass et al., 2013). Bailey et al. (2014) suggest that success on math games may positively alter children's beliefs about the value of math and required effort, thereby enhancing long-term math achievement (Blackwell, Trzesniewski, & Dweck, 2007). In addition, peer-tutoring programs have reciprocal benefits, in

that they effectively increase fluency for both the tutee (i.e., through overt responding) and the tutor (i.e., by delivering feedback; Rhymer, Dittmer, Skinner, & Jackson, 2000).

Finally, based on the results of my study, interventions provided over the summer months may help inattentive children at risk of math failure catch up to their peers. Of particular interest are programs such as the Summer Treatment Program (STP; e.g., Fabiano, Schatz, & Pelham, 2014; Pelham et al., 2000), a multimodal intensive treatment program based on behavioral principles (operant conditioning and learning theory), which targets multiple areas of impairment typically seen in children with ADHD (i.e., academic functioning, family and peer relationships). Academic impairments are addressed through daily classroom instruction, in an environment that promotes the development of adaptive skills and academic enablers (e.g., cooperation, increasing tolerance for seat work). Although STP has empirical support demonstrating its effectiveness for children with ADHD (see Fabiano et al., 2014, for a review), it is currently unclear whether a similar approach (e.g., with differing intensity) would be useful for inattentive children at risk of math difficulties, but whose behavioral symptoms do meet clinical threshold. Future studies would be needed to explore these possibilities.

**Limitations and Future Directions**

Although the current findings are notable, it is important to highlight the study's limitations. First, although this was a community-based study, the ethnic composition of the study was largely Caucasian. Previous studies examining the relationship between ethnicity, attention problems, and academic achievement suggests a potential relationship between these factors. For example, Hooper and colleagues (2010) found that ratings of inattention predicted a more profound impact on academic functioning for African-American compared to Caucasian students (Hooper, Roberts, Sideris, Burchinal, & Zeisel, 2010). Further, Rabiner and colleagues (2004) found that attention problems were uniquely associated with academic achievement in

first grade children, and that a substantial portion of the achievement gap between African Americans and Caucasian students was related to higher rates of inattention in the former group. Although these studies were conducted using American samples, and it is unclear if similar findings would be seen using a Canadian sample; nonetheless, with potential interactions between these factors, replication with large samples representative of the population in terms of ethnic representation will be important. Second, the significant amount of variance that remained unexplained in each analysis  of the current study highlights the importance of including of other variables that are associated with math fluency in future longitudinal work (Bailey et al., 2014). The current study only accounted for a limited number of domain-general abilities. However, previous studies as well as extant theory in math development suggest important roles for both domain-general and domain- specific factors. For example, research highlights the importance of domain-general constructs such as IQ (Geary, 2011a), processing speed (Bull & Johnston, 1997), and phonological processing (Barnes et al., 2014; De Smedt et al., 2010; Fuchs et al., 2005; Fuchs et al., 2006) for fluency development (but also see Jordan, Hanich, and Kaplan; 2003). Domain-specific abilities (e.g., number sense and subitizing ability) are also important in math fluency development (Fuchs et al., 2010; Martin et al., 2014; Robinson, Menchetti, & Torgesen, 2002).

Other important factors may include academic enablers, such as motivation or task engagement (Plamondon & Martinussen, 2015; Murayama et al., 2013). In terms of theories proposed to synthesize current findings on math development, LeFevre and colleagues' Pathways to Mathematics model (LeFevre et al., 2010; Sowinski et al., 2015) highlights the contribution of linguistic and quantitative pathways, in addition to a visual attention (working memory) pathway, to children's skill in numeration, number line understanding, and calculation (LeFevre et al., 2010). Notably, all three pathways contributed to math fluency specifically, while controlling for sex, parent level of education, and processing speed (Sowinski et al., 2015).

Therefore, linguistic and quantitative factors should be included in future longitudinal studies on math fluency development. Further, an appreciation of the fact that the independent variables may themselves be developing (e.g., working memory; Li and Geary, 2013) will be important to include in these studies. Third, providing a direct comparison of addition and subtraction fluency using the corresponding problem stems (e.g., 7 + 4 and 7 - 4), would have allowed for a more robust comparison of growth trajectories (Barrouillet et al., 2008; Kamii et al., 2001). Fourth, a small number of longitudinal studies on the math development have noted heterogeneity in terms of the developmental dynamics of math achievement in community samples (Aunola et al., 2004; Jordan et al., 2009; Jordan et al., 2006; Salaschek et al., 2014). Although the current study was able to model variability in the sample level, data-driven analyses identifying subgroups of children who follow different trajectories, and the individual variables that predict these trajectories in community samples (Salaschek et al., 2014) are important next steps, which would further differentiate children who are particularly at risk for long-term math failure versus those who may be able to catch up to their peers, despite early struggles in math.

Finally, the current research did not take into account specific teaching practices that were in operation at the time of the study, and the potential effects of these practices on math fluency development for the current sample. Although it might be assumed that approaches to teaching may be fairly uniform (i.e., following the Ontario curriculum), there may be variability in the implementation of prescribed methods (Ross et al., 2002). Factors to consider would include whether students were offered instruction aimed at enhancing fluency skill, whether a focus on fluency shifted over the grades, and whether practices were differentiated according to individual student need. These variables could have important implications for results both between and within students over time (e.g., could influence the Matthew effect). This type of analysis could be performed in future research by adding a third level to the HLM analysis,

which would account for classroom variables (i.e., time is nested within students, and students are nested within classrooms).

**Summary and Conclusion**

The current study showed that basic addition and subtraction fluency develops in a cumulative manner, both within and across the early elementary grades. Thus, children who lag behind their peers in fluency skills do not catch up, at least in the four-year period covered by this study. Further, by parceling out seasonal effects, I showed that visual-spatial working memory is related to the development of fluency during instructional periods (i.e., school), whereas classroom inattention was a unique predictor of weaker addition and subtraction fluency both during school and over the summer. Further, teacher ratings of classroom inattention predicted children's long-term fluency development and the widening gap between students across the early elementary years, even after controlling for age, sex, parent level of education, verbal and visual-spatial working memory. Major implications of these findings include a potential role for behavioral symptoms of classroom inattention as a flag for later math difficulties, as well as the consideration of the interaction between timing of interventions and individual characteristics when planning for effective math fluency instruction and intervention. These results also suggest that children who are inattentive and have weak visual-spatial working memory may be particularly at risk for math difficulties in the early elementary grades.

The results of the current dissertation raise interesting questions for future research. Longitudinal studies using larger and more representative samples are needed that also incorporate piecewise designs to explore the interaction between instructional breaks and individual characteristics (i.e., demographic, cognitive, and behavioral). Further, studies that examine whether inattention in the classroom is a reliable marker for mathematical difficulties, those that explore the reciprocal impact of interventions on both fluency development and

behavioral attention, and those that evaluate strategies that could mitigate the negative impact of inattentive behavior on mathematical achievement while increasing productivity are indicated. Finally, the finding inattention and working memory also shared variance in the prediction of the development of math fluency across the grades highlights a close relationship between these factors. Examination of this association, with a focus on potential reciprocal relationships (e.g., Metcalfe et al., 2013), and exploration of the value of shared (as opposed to unique) variance (Plamondon &Martinussen, 2015) is warranted. The overarching goal of this line of research would be to gain a clearer understanding of the relationship between teacher-related inattention, working memory, and math achievement, to guide the development of preventative instructional or early-intervention programs geared towards helping children achieve their academic potential.

Table 1. *Participant demographics at study entry (fall of year 1).*

|  | Grade at Study Entry | | |
|---|---|---|---|
|  | Grade 1 | Grade 2 | Grade 3 |
| n | 55 | 66 | 88 |
| Sex (% male) | 45% | 48% | 54% |
| Mean age in years | 6.05 | 7.04 | 8.04 |
|  | Exceptionalities (Percentages) | | |
| Gifted | 0 | 0 | 0 |
| Learning Disability | 3.6 | 1.5 | 3.6 |
| Language Impairment | 7.3 | 6.1 | 1.2 |
| Mild Intellectual Disability | 0 | 0 | 0 |
| Autism/ASD | 0 | 0 | 0 |
| Multiple Exceptionalities | 0 | 0 | 0 |
| Developmental Disability | 0 | 1.5 | 0 |
| Behavioural | 3.6 | 0 | 1.2 |
| Attention Deficit Hyperactivity Disorder | 1.8 | 4.5 | 7.2 |

Table 2. *Accelerated growth curve design.*

| Grade at study entry (cohort) | Grade 1 Fall | Grade 1 Spring | Grade 2 Fall | Grade 2 Spring | Grade 3 Fall | Grade 3 Spring | Grade 4 Fall | Grade 4 Spring |
|:---:|:---:|:---:|:---:|:---:|:---:|:---:|:---:|:---:|
| 1 | X | X | X | X | | | | |
| 2 | | | X | X | X | X | | |
| 3 | | | | | X | X | X | X |

*Note. X*s indicate points where students were assessed.

Table 3. *Contrasts depicting linear, quadratic, and cubic trends, for both between-grade and within-grade growth.*

| Grade (Time) | Between-grade linear | Between-grade quadratic | Between-grade cubic | Within-grade mean | Within-grade linear | Within Grade Quad |
|---|---|---|---|---|---|---|
| 1 (Time 1) | -3 | 1 | -1 | -1 | 3 | -1 |
| 1 (Time 2) | -3 | 1 | -1 | 1 | -3 | 1 |
| 2 (Time 1) | -1 | -1 | 3 | -1 | 1 | 1 |
| 2 (Time 2) | -1 | -1 | 3 | 1 | -1 | -1 |
| 3 (Time 1) | 1 | -1 | -3 | -1 | -1 | 1 |
| 3 (Time 2) | 1 | -1 | -3 | 1 | 1 | -1 |
| 4 (Time 1) | 3 | 1 | 1 | -1 | -3 | -1 |
| 4 (Time 2) | 3 | 1 | 1 | 1 | 3 | 1 |

Table 4. *Means and Standard Deviations of Math Fluency.*

| | Grade at study entry (cohort) | Grade 1 Time 1 | Grade 1 Time 2 | Grade 2 Time 1 | Grade 2 Time 2 | Grade 3 Time 1 | Grade 3 Time 2 | Grade 4 Time 1 | Grade 4 Time 2 |
|---|---|---|---|---|---|---|---|---|---|
| CBM Addition | 1 (n = 52) | 5.00 (4.18) | 8.85 (4.82) | 9.91 (6.36) | 15.63 (8.30) | | | | |
| | 2 (n = 61) | | | 11.83 (5.79) | 15.53 (6.31) | 17.15 (5.90) | 21.63 (7.83) | | |
| | 3 (n = 79) | | | | | 15.61 (6.27) | 21.12 (6.85) | 19.94 (7.79) | 22.96 (9.48) |
| CBM Subtraction | 1 (n = 53) | 3.12 (2.87) | 6.13 (3.56) | 7.72 (4.10) | 11.00 (5.48) | | | | |
| | 2 (n = 61) | | | 6.67 (4.43) | 10.06 (4.98) | 12.27 (5.01) | 14.32 (5.75) | | |
| | 3 (n = 79) | | | | | 10.79 (4.81) | 13.83 (5.44) | 14.32 (5.80) | 17.18 (6.72) |
| WJ-III Math Fluency | 1 (n = 53) | 3.61 (2.89) | 5.73 (344) | 7.54 (3.65) | 9.75 (3.91) | | | | |
| | 2 (n = 62) | | | 7.92 (3.79) | 9.07 (3.83) | 10.89 (4.20) | 13.06 (4.28) | | |
| | 3 (n = 79) | | | | | 11.26 (3.58) | 12.00 (3.87) | 13.66 (4.30) | 15.02 (5.17) |

*Note.* Numbers in parentheses are standard deviations.

Table 5. *Correlations Among Predictor Variables and Addition Fluency (n =192 to 201).*

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| 1. CBM Addition T1 | - | - | - | - | - | - | - | - |
| 2. CBM Addition T2 | .84*** | - | - | - | - | - | - | - |
| 3. CBM Addition T3 | .79*** | .81*** | - | - | - | - | - | - |
| 4. CBM Addition T4 | .72*** | .72*** | .81*** | - | - | - | - | - |
| 5. Sex | -.09 | -.07 | -.07 | -.06 | - | - | - | - |
| 6. Parent Education | .07 | .11 | .14 | .22** | .08 | - | - | - |
| 7. Verbal WM | .17* | .20** | .21** | .21** | -.07 | .17* | - | - |
| 8. Visual-Spatial WM | .38*** | .50*** | .41*** | .45*** | .05 | .15* | .14* | - |
| 9. Inattention | -.32*** | -.35*** | -.43*** | -.48*** | -.38*** | -.23** | -.16* | -.36*** |

*Note.* [a] Point biserial correlations are reported for Sex.

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

Table 6. *Correlations Among Predictor Variables and Subtraction Fluency (n =193 to 203)*

|  | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| 1. CBM Subtraction T1 | - | - | - | - | - | - | - | - |
| 2. CBM Subtraction T2 | .78*** | - | - | - | - | - | - | - |
| 3. CBM Subtraction T3 | .73*** | .81*** | - | - | - | - | - | - |
| 4. CBM Subtraction T4 | .68*** | .76*** | .80*** | - | - | - | - | - |
| 5. Sex | -.15* | -.11 | -.04 | -.19** | - | - | - | - |
| 6. Parent Education | .05 | .09 | .17* | .19** | .09 | - | - | - |
| 7. Verbal WM | .16* | .17* | .19** | .22** | -.06 | .17* | - | - |
| 8. Visual-Spatial WM | .41*** | .49*** | .43*** | .47*** | .04 | .15* | .12 | - |
| 9. Inattention | -.29*** | -.35*** | -.45*** | -.38*** | -.39*** | -.24** | -.18** | -.34*** |

*Note.* [a] Point biserial correlations are reported for Sex.

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

Table 7. *Correlations Among Predictor Variables and WJ-III Fluency (n =193 to 203)*

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| 1. WJ-III T1 | - | - | - | - | - | - | - | - |
| 2. WJ-III T2 | .85*** | - | - | - | - | - | - | - |
| 3. WJ-III T3 | .84*** | .88*** | - | - | - | - | - | - |
| 4. WJ-III T4 | .77*** | .82*** | .83*** | - | - | - | - | - |
| 5. Sex | -.07 | -.09 | -.09 | -.12 | - | - | - | - |
| 6. Parent Education | .11 | .09 | .13 | .16* | .09 | - | - | - |
| 7. Verbal WM | .21** | .23** | .24** | .24** | -.06 | .17* | - | - |
| 8. Visual-Spatial WM | .45*** | .47*** | .42*** | .41*** | .04 | .15* | .12 | - |
| 9. Inattention | -.35*** | -.41*** | -.42*** | -.40*** | -.39*** | -.24** | -.18** | -.34*** |

*Note.* [a] Point biserial correlations are reported for Sex.
*p < .05, **p <. 01, *** p < .001

Table 8. *Addition Growth Models with Predictors Entered at Level-2 (Level-1 n = 785, Level-2 n = 200)*

| | FUM | Model 0 | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 | Model 7 |
|---|---|---|---|---|---|---|---|---|---|
| **For Intercept, $\beta_0$** | | | | | | | | | |
| Intercept, $\gamma_{00}$ | 15.95*** | 15.14*** | 15.05*** | 15.38*** | 15.54*** | 15.38*** | 15.43*** | 16.84*** | 14.99*** |
| Sex, $\gamma_{01}$ | | | | -.63 | -.93 | -.65 | -.89 | -3.76*** | - |
| Parent Education, $\gamma_{02}$ | | | | | .91** | .73** | .54* | .17 | .15 |
| Verbal WM, $\gamma_{03}$ | | | | | | .54*** | .45** | .25* | .36** |
| Visual-spatial WM, $\gamma_{04}$ | | | | | | | .50*** | .18* | .23* |
| Inattention, $\gamma_{05}$ | | | | | | | | -2.76*** | -2.17*** |
| **For Between Gr. Linear, $\beta_1$** | | | | | | | | | |
| Intercept, $\gamma_{10}$ | | 2.38*** | 2.38*** | 2.41*** | 2.46*** | 2.44*** | 2.44*** | 2.72*** | 2.45*** |
| Sex, $\gamma_{11}$ | | | | -.05 | -.13 | -.10 | -.16 | -.54* | - |
| Parent Education, $\gamma_{12}$ | | | | | .19 | .17 | .13 | .08 | .05 |
| Verbal WM, $\gamma_{13}$ | | | | | | .05 | .03 | .03 | .04 |
| Visual-spatial WM, $\gamma_{14}$ | | | | | | | .09*** | .05 | .06 |
| Inattention, $\gamma_{15}$ | | | | | | | | -.28** | -.20* |
| **For Between Gr. Quad, $\beta_2$** | | | | | | | | | |
| Intercept, $\gamma_{20}$ | | -.72*** | -.77*** | -.87*** | -.85** | -.83** | -.86** | -.73** | -.80*** |
| Sex, $\gamma_{21}$ | | | | .21 | .19 | .13 | .11 | -.17 | - |
| Parent Education, $\gamma_{22}$ | | | | | -.01 | .02 | .02 | -.03 | -.02 |
| Verbal WM, $\gamma_{23}$ | | | | | | -.12* | -.11 | -.14* | -.13* |
| Visual-spatial WM, $\gamma_{24}$ | | | | | | | -.03 | -.07 | -.06 |
| Inattention, $\gamma_{25}$ | | | | | | | | -.31* | -.28* |
| **For Within Gr. Mean, $\beta_3$** | | | | | | | | | |
| Intercept, $\gamma_{30}$ | | 2.05*** | 2.05*** | .2.07*** | 2.09*** | 2.09** | 2.13*** | 2.20*** | 2.07*** |
| Sex, $\gamma_{31}$ | | | | -.04 | -.08 | -.07 | -.13 | -.27 | - |
| Parent Education, $\gamma_{32}$ | | | | | .19* | .18* | .13 | .11 | .11 |
| Verbal WM, $\gamma_{33}$ | | | | | | .02 | -.00 | -.00 | -.00 |
| Visual-spatial WM, $\gamma_{34}$ | | | | | | | .14*** | .13*** | .13*** |
| Inattention, $\gamma_{35}$ | | | | | | | | -.13 | -.09 |
| **For Within Gr. Linear, $\beta_4$** | | | | | | | | | |
| Intercept, $\gamma_{40}$ | | -.03 | -.05 | -.05 | -.04 | -.05 | -.10 | -.10 | -.10 |
| **For Within Gr. Quad, $\beta_5$** | | | | | | | | | |
| Intercept, $\gamma_{50}$ | | -.35** | -.32** | -.20 | -.21 | -.210 | -.19 | -.18 | -.31* |
| Sex, $\gamma_{51}$ | | | | -.24 | -.23 | -.24 | -.24 | -.27 | - |
| Parent Education, $\gamma_{52}$ | | | | | -.00 | .00 | -.00 | -.00 | -.00 |
| Verbal WM, $\gamma_{53}$ | | | | | | -.03 | -.02 | -.02 | -.01 |
| Visual-spatial WM, $\gamma_{54}$ | | | | | | | .01 | .01 | .01 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Inattention, $\gamma_{55}$ | | | | | | | | -.02 | .01 |

**Random effect**

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Intercept, $u_0$ | 48.01*** | 33.79*** | 33.45*** | 33.21*** | 31.40*** | 28.96*** | 25.07*** | 15.77*** | 18.46*** |
| Between Grade Linear, $u_{01}$ | | | .95*** | .95*** | .88*** | .86*** | .72*** | .50*** | .57*** |
| Within Grade Mean, $u_{02}$ | | | .43 | .42 | .36 | .35 | .18 | .14 | .16 |
| Level-1 error $(r_{ij})$[1] | 27.23 | 13.94 | 11.65 | 11.52 | 11.48 | 11.34 | 11.24 | 11.43 | 11.43 |
| Total % Variance intercept explained | | | | .07 | 6.12 | 13.42 | 25.05 | 52.85 | 44.81 |
| Total % Variance linear slope explained | | | | .00 | 7.36 | 9.47 | 24.21 | 47.36 | 40.00 |
| Model fit deviance test compared to previous model | | | $\chi^2 (3) =$ 21.72*** | $\chi^2 (5) =$ 2.37 | $\chi^2 (5) =$ 12.42* | $\chi^2 (5) =$ 19.22** | $\chi^2 (5) =$ 33.65*** | $\chi^2 (5) =$ 87.40*** | $\chi^2 (5) =$ 28.72*** |

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

[1]*Note*: $p$ values are not calculated for Level 1 error. FUM: Fully Unconditional Model. Model 0 represents a preliminary unconditional model (no predictors at level 2) where only the intercept was treated as random; all other contrasts were fixed. Model 1 represents is the final unconditional model (no predictors at level 2), when between-grade linear and within grade means were allowed to vary (best fit as outlined in the text). Model 2 through 6 represents forward stepping of variables in the following order: sex, parent level of education, verbal working memory, visual-spatial working memory, classroom inattention. Backward stepping was used in Model 7, where the effect of sex was removed.

Table 9. *Subtraction Growth Models with Predictors Entered at Level-2 (Level-1 n = 789, Level 2 n= 201)*

| Fixed effect | FUM | Model 0 | Model 1 | Model 2 | Model 3 | Model 4 | Model5 | Model 6 | Model 7 |
|---|---|---|---|---|---|---|---|---|---|
| **For Intercept, $\beta_0$** | | | | | | | | | |
| Intercept, $\gamma_{00}$ | 10.97*** | 10.51*** | 10.61*** | 11.11*** | 11.24*** | 11.14*** | 11.20*** | 12.18*** | 10.60*** |
| Sex, $\gamma_{01}$ | | | | -.97 | -1.22* | -1.03 | -1.17* | -3.17*** | - |
| Parent Education, $\gamma_{02}$ | | | | | .67** | .53** | .39* | .15 | .12 |
| Verbal WM, $\gamma_{03}$ | | | | | | .41*** | .35*** | .20* | .29*** |
| Visual-spatial WM, $\gamma_{04}$ | | | | | | | .38*** | .18** | .22** |
| Inattention, $\gamma_{05}$ | | | | | | | | -1.83*** | -1.34*** |
| **For Between Gr. Linear, $\beta_1$** | | | | | | | | | |
| Intercept, $\gamma_{10}$ | | 2.01*** | 2.04*** | 2.05*** | 2.09*** | 2.08*** | 2.04*** | 2.20*** | 2.00*** |
| Sex, $\gamma_{11}$ | | | | -.04 | -.11 | -.08 | -.13 | -.43** | - |
| Parent Education, $\gamma_{12}$ | | | | | .18** | .15* | .12 | .08 | .06 |
| Verbal WM, $\gamma_{13}$ | | | | | | .04 | .03 | .02 | .04 |
| Visual-spatial WM, $\gamma_{14}$ | | | | | | | .04* | .02 | .02 |
| Inattention, $\gamma_{15}$ | | | | | | | | -.21*** | -.16** |
| **For Within Gr. Mean, $\beta_3$** | | | | | | | | | |
| Intercept, $\gamma_{20}$ | | 1.45*** | 1.44*** | 1.64*** | 1.66*** | 1.65*** | 1.67*** | 1.72*** | 1.44*** |
| Sex, $\gamma_{21}$ | | | | -.40* | -.43* | -.41* | -.44* | -.54* | - |
| Parent Education, $\gamma_{22}$ | | | | | .09 | .08 | .05 | .04 | .03 |
| Verbal WM, $\gamma_{23}$ | | | | | | .03 | .02 | .01 | .02 |
| Visual-spatial WM, $\gamma_{24}$ | | | | | | | .08*** | .07** | .08** |
| Inattention, $\gamma_{25}$ | | | | | | | | -.09 | -.00 |
| **Random Effect** | | | | | | | | | |
| Intercept, $u_0$ | 24.17*** | 17.75*** | 16.87*** | 16.64*** | 15.79*** | 14.65*** | 12.54*** | 7.88*** | 9.95*** |
| Between Grade Linear, $u_{01}$ | | | .49*** | .49*** | .43*** | .43*** | .36*** | .22*** | .29*** |
| Within Grade Mean, $u_{02}$ | | | .31 | .29 | .27 | .27 | .19 | .14 | .22 |
| Level-1 error $(r_{ij})$[1] | 16.85 | 8.33 | 7.10 | 7.06 | 7.03 | 7.02 | 7.03 | 7.22 | 7.13 |
| Total % Variance intercept explained | | | | 1.36 | 6.40 | 13.15 | 25.66 | 53.28 | 41.01 |
| Total % Variance linear slope explained | | | | .00 | 12.24 | 12.24 | 26.53 | 55.10 | 40.81 |
| Model fit deviance test compared to previous model | | | $\chi^2(3) =$ 21.39*** | $\chi^2(3) =$ 5.44 | $\chi^2(3) =$ 13.29** | $\chi^2(3) =$ 13.05** | $\chi^2(3) =$ 28.99*** | $\chi^2(3) =$ 78.02*** | $\chi^2(3) =$ 38.49*** |

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

*Note*: *p* values are not calculated for Level 1 error. FUM: Fully Unconditional Model. Model 0 represents a preliminary unconditional model (no predictors at level 2) where only the intercept was treated as random; all other contrasts were fixed. Model 1 represents is the final unconditional model (no predictors at level 2), when between-grade linear and within grade means were allowed to vary (best fit as outlined in the text). Model 2 through 6 represents forward stepping of variables in the following order: sex, parent level of education, verbal working memory, visual-spatial working memory, classroom inattention. Backward stepping was used in Model 7, where the effect of sex was removed.

Table 10. *WJ-III Growth Models with Predictors Entered at Level-2 (Level-1 n = 790, Level-2 n= 201)*
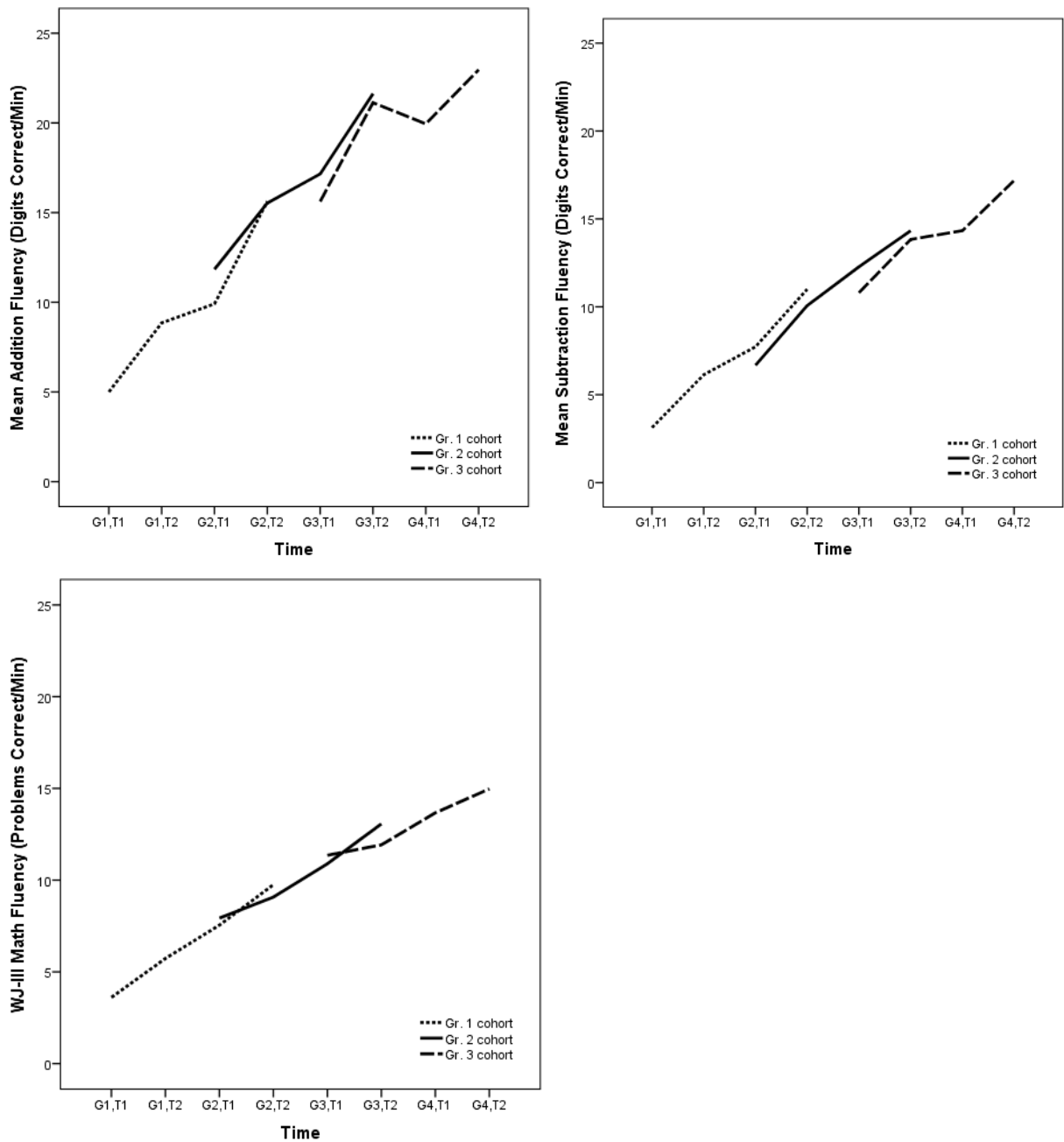
| Fixed effect | FUM | Model 0 | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 | Model 7 |
|---|---|---|---|---|---|---|---|---|---|
| For Intercept, $\beta_0$ | | | | | | | | | |
| Intercept, $\gamma_{00}$ | 10.27*** | 9.81*** | 9.89*** | 10.16*** | 10.24*** | 10.14*** | 10.21*** | 11.10*** | 9.91*** |
| Sex, $\gamma_{01}$ | | | | -.55 | -.70 | -.52 | -.61 | -2.41*** | - |
| Parent Education, $\gamma_{02}$ | | | | | .51** | .38* | .28 | .05 | .04 |
| Verbal WM, $\gamma_{03}$ | | | | | | .41*** | .37*** | .23*** | .29*** |
| Visual-spatial WM, $\gamma_{04}$ | | | | | | | .30*** | .11* | .14** |
| Inattention, $\gamma_{04}$ | | | | | | | | -1.68*** | -1.31*** |
| For Between Grade Linear, $\beta_1$ | | | | | | | | | |
| Intercept, $\gamma_{10}$ | | 1.63*** | 1.64*** | 1.68*** | 1.70*** | 1.69*** | 1.66*** | 1.74*** | 1.64*** |
| Sex, $\gamma_{11}$ | | | | -.10 | -.12 | -.11 | -.13 | -.21 | - |
| Parent Education, $\gamma_{12}$ | | | | | .05 | .04 | .02 | .02 | .00 |
| Verbal WM, $\gamma_{13}$ | | | | | | .02 | .02 | .02 | .02 |
| Visual-spatial WM, $\gamma_{14}$ | | | | | | | .01 | .00 | .00 |
| Inattention, $\gamma_{15}$ | | | | | | | | -.05 | -.02 |
| For Between Gr. Quad, $\beta_2$ | | | | | | | | | |
| Intercept, $\gamma_{20}$ | | -.29*** | -.23** | -.16 | -.16 | -.15 | -.14 | -.08 | -.21* |
| Sex, $\gamma_{21}$ | | | | -.16 | -.15 | -.16 | -.16 | -.28 | - |
| Parent Education, $\gamma_{22}$ | | | | | -.01 | -.00 | -.00 | -.03 | -.02 |
| Verbal WM, $\gamma_{23}$ | | | | | | -.01 | -.01 | -.03 | -.02 |
| Visual-spatial WM, $\gamma_{24}$ | | | | | | | .00 | -.01 | -.01 |
| Inattention, $\gamma_{25}$ | | | | | | | | -.14* | -.10 |
| For Within Grade Mean, $\beta_3$ | | | | | | | | | |
| Intercept, $\gamma_{30}$ | | .77*** | .76*** | .82*** | .82*** | .82*** | .82*** | .86*** | .76*** |
| Sex, $\gamma_{31}$ | | | | -.12 | -.12 | -.11 | -.12 | -.19 | - |
| Parent Education, $\gamma_{32}$ | | | | | .00 | -.00 | -.00 | -.00 | -.00 |
| Verbal WM, $\gamma_{34}$ | | | | | | .00 | .00 | .00 | .00 |
| Visual-spatial WM, $\gamma_{35}$ | | | | | | | .00 | -.00 | -.00 |
| Inattention, $\gamma_{36}$ | | | | | | | | -.06 | -.03 |
| Random Effect | | | | | | | | | |
| Intercept, $u_0$ | 18.09*** | 12.10*** | 11.54*** | 11.51*** | 10.93*** | 9.69*** | 8.51*** | 4.91*** | 5.91*** |
| Between Grade Linear, $u_{01}$ | | | .23** | .22** | .22* | .22* | .20* | .20* | .24** |
| Within Grade Mean, $u_{02}$ | | | .20* | .19* | .19* | .19* | .18* | .20* | .22* |
| Level-1 error $(r_{ij})$[1] | 7.95 | 3.07 | 3.07 | 3.08 | 3.07 | 3.07 | 3.09 | 3.05 | 3.07 |

| | | | | | | |
|---|---|---|---|---|---|---|
| % Variance intercept explained | .25 | 5.28 | 16.03 | 26.25 | 57.45 | 48.78 |
| % Variance linear slope explained | 4.34 | 4.34 | 4.34 | 13.04 | 13.04 | -4.16 |
| Model fit deviance test compared to previous model | $\chi^2 (3) =$ 16.81*** | $\chi^2 (4) =$ 2.52 | $\chi^2 (4) =$ 9.90* | $\chi^2 (4) =$ 22.45*** | $\chi^2 (4) =$ 24.11*** | $\chi^2 (4) =$ 102.69*** | $\chi^2 (4) =$ 35.56*** |

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

[1]*Note*: $p$ values are not calculated for Level 1 error. FUM: Fully Unconditional Model. Model 0 represents a preliminary unconditional model (no predictors at level 2) where only the intercept was treated as random; all other contrasts were fixed. Model 1 represents is the final unconditional model (no predictors at level 2), when between-grade linear and within grade means were allowed to vary (best fit as outlined in the text). Model 2 through 6 represents forward stepping of variables in the following order: sex, parent level of education, verbal working memory, visual-spatial working memory, classroom inattention. Backward stepping was used in Model 7, where the effect of sex was removed.

*Figure 1.* Accelerated growth curve for CBM addition (n = 192), CBM subtraction (n = 193), and WJ-III fact fluency (n = 194). Fluency scores represent raw score means. In each graph, the three distinct curves represent the three grade cohorts (Grade 1-3, at study entry).
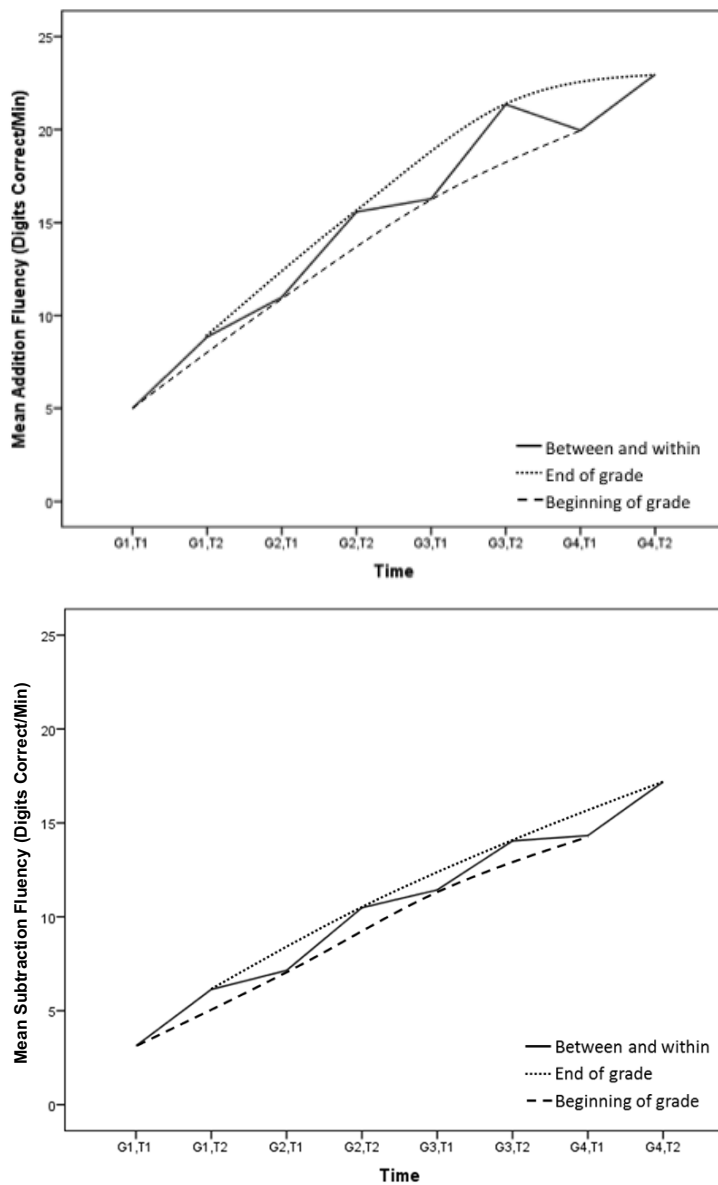
*Figure 2.* Addition (2a) and subtraction (2b) fluency growth curves. For both graphs, solid black lines depict the combined effect of the between- and within-grade growth. In Figure 2a, significant between-grade linear and between-grade quadratic growth trends are illustrated by both the dotted (linking the beginning of the year of each grade) and hashed lines (linking the end of year for each grade). Similarly, Figure 2b, shows the significant between-grade linear growth.
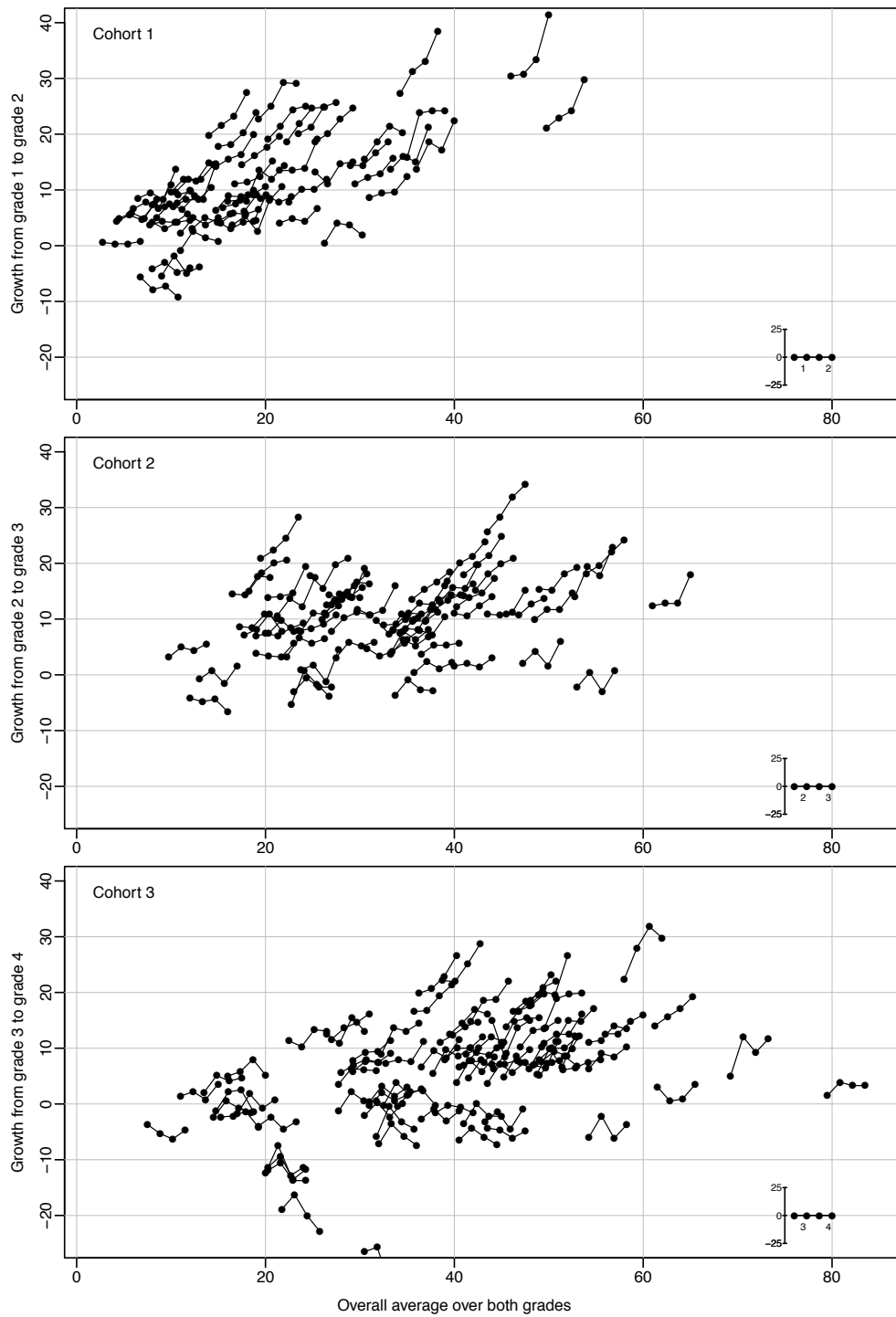
*Figure 3a.* Plot of individual growth curves representing CBM addition. Individual curves are plotted according to the average growth over four testing points (X-axis) by growth from Year 1 to Year 2 of the study (Y-axis), for each grade at study entry (i.e., Grade 1, 2, 3 at study entry).
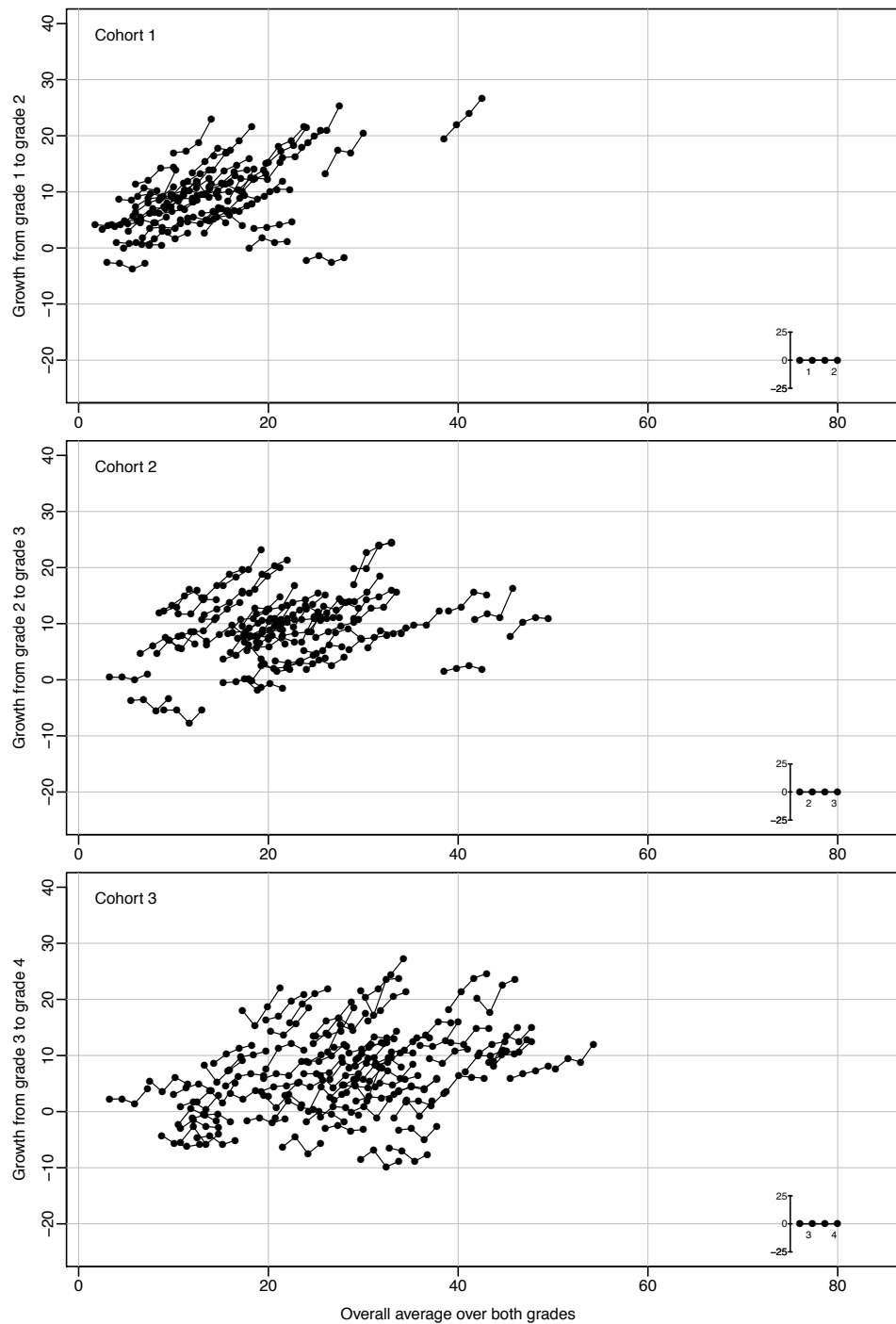
*Figure 3b.* Plot of individual growth curves representing CBM subtraction. Individual curves are

plotted according to the average growth over four testing points (X-axis) by growth from Year 1 to

Year 2 of the study (Y-axis), for each grade at study entry (i.e., Grade 1, 2, 3 at study entry).
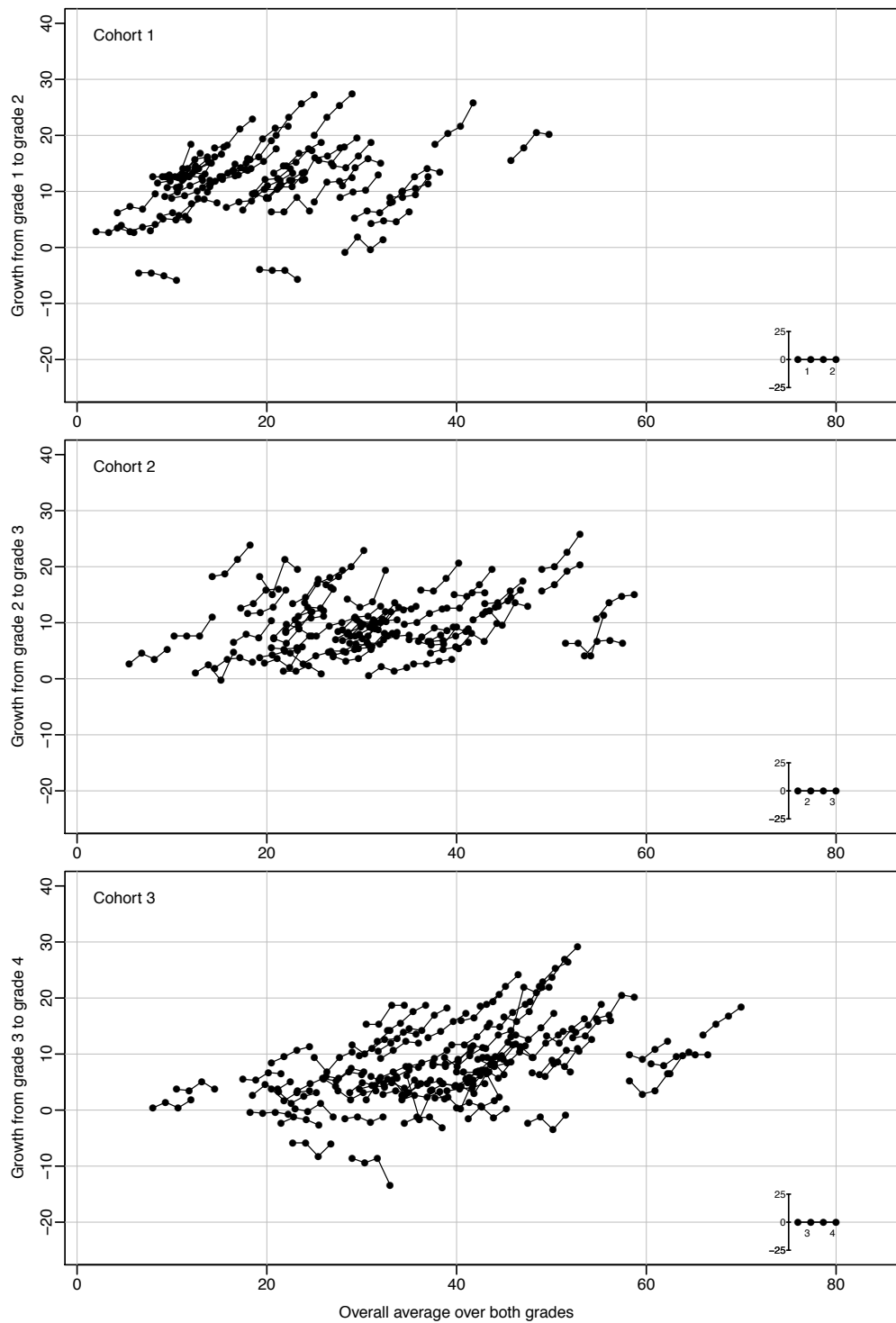
*Figure 3c.* Plot of individual growth curves representing WJ-III. Individual curves are plotted

according to the average growth over four testing points (X-axis) by growth from Year 1 to Year 2 of

the study (Y-axis), for each grade at study entry (i.e., Grade 1, 2, 3 at study entry).
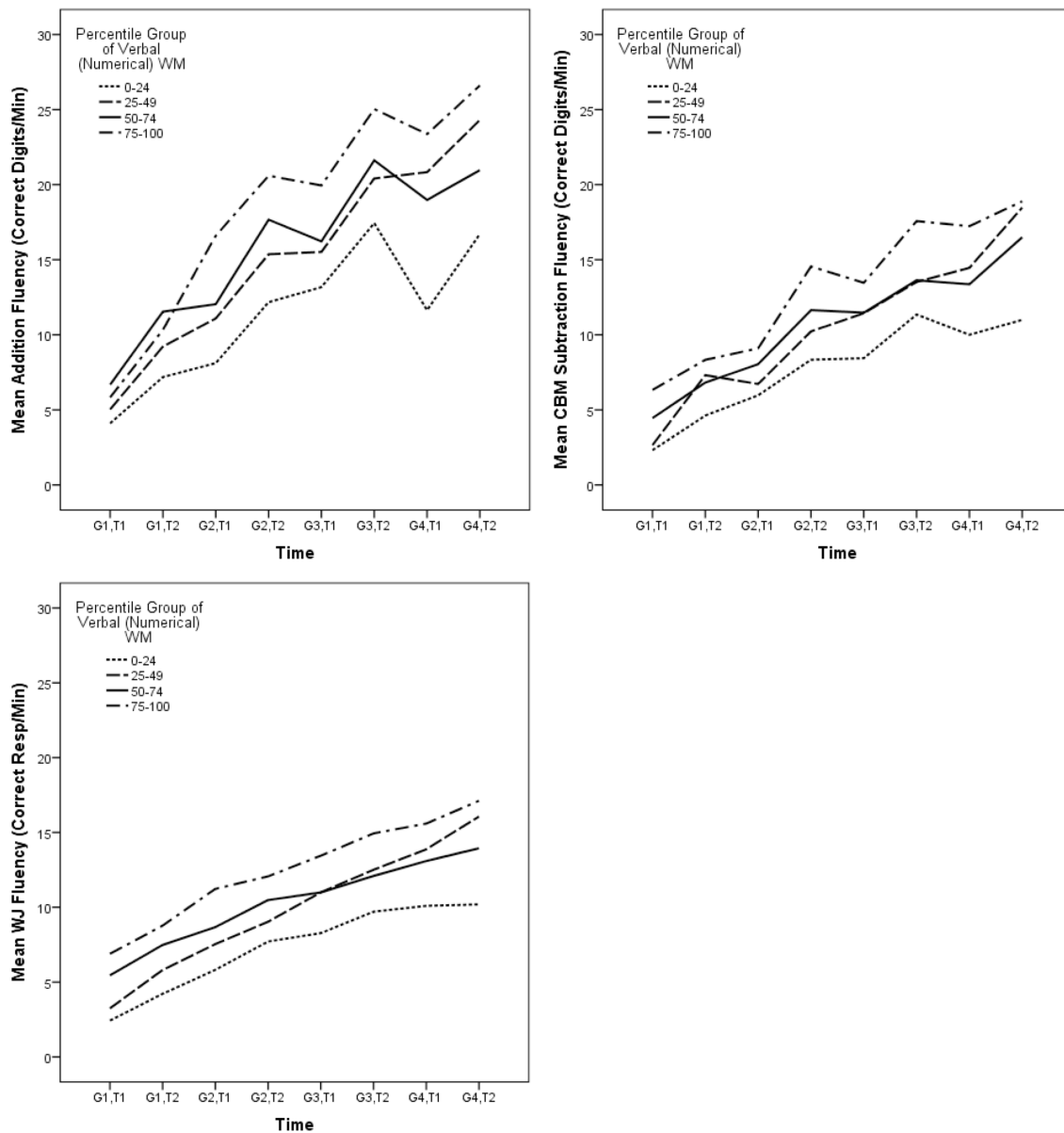
*Figure 4.* Relationship between verbal working memory and math fluency for a) CBM addition, b) CBM subtraction, and c) WJ-III Math Fluency. Lines depict quartiles of verbal working memory.

*Figure 5.* Relationship between visual-spatial working memory and math fluency for a) CBM addition, b) CBM subtraction, and c) WJ-III Math Fluency. Lines depict quartiles of visual-spatial working memory.
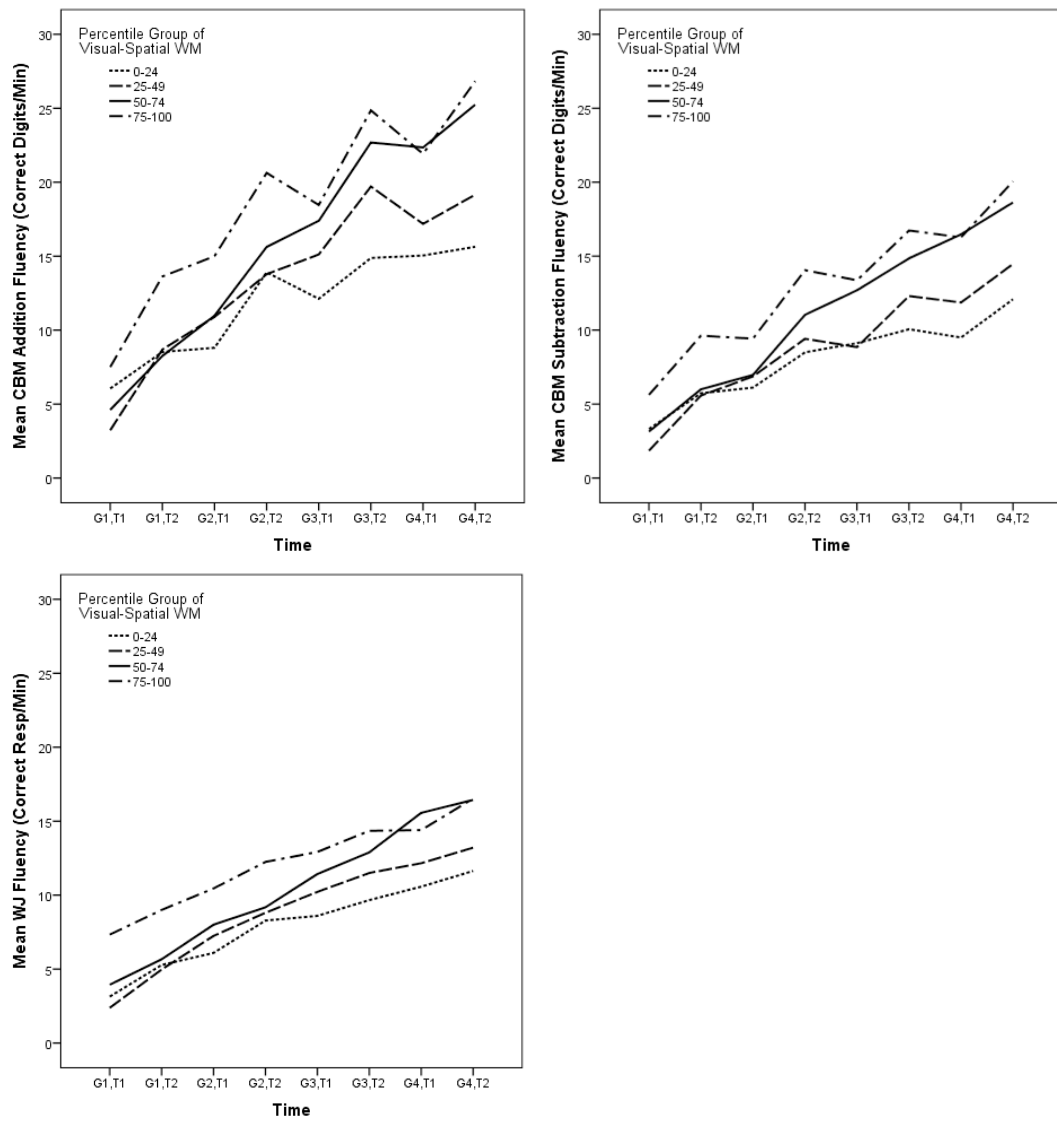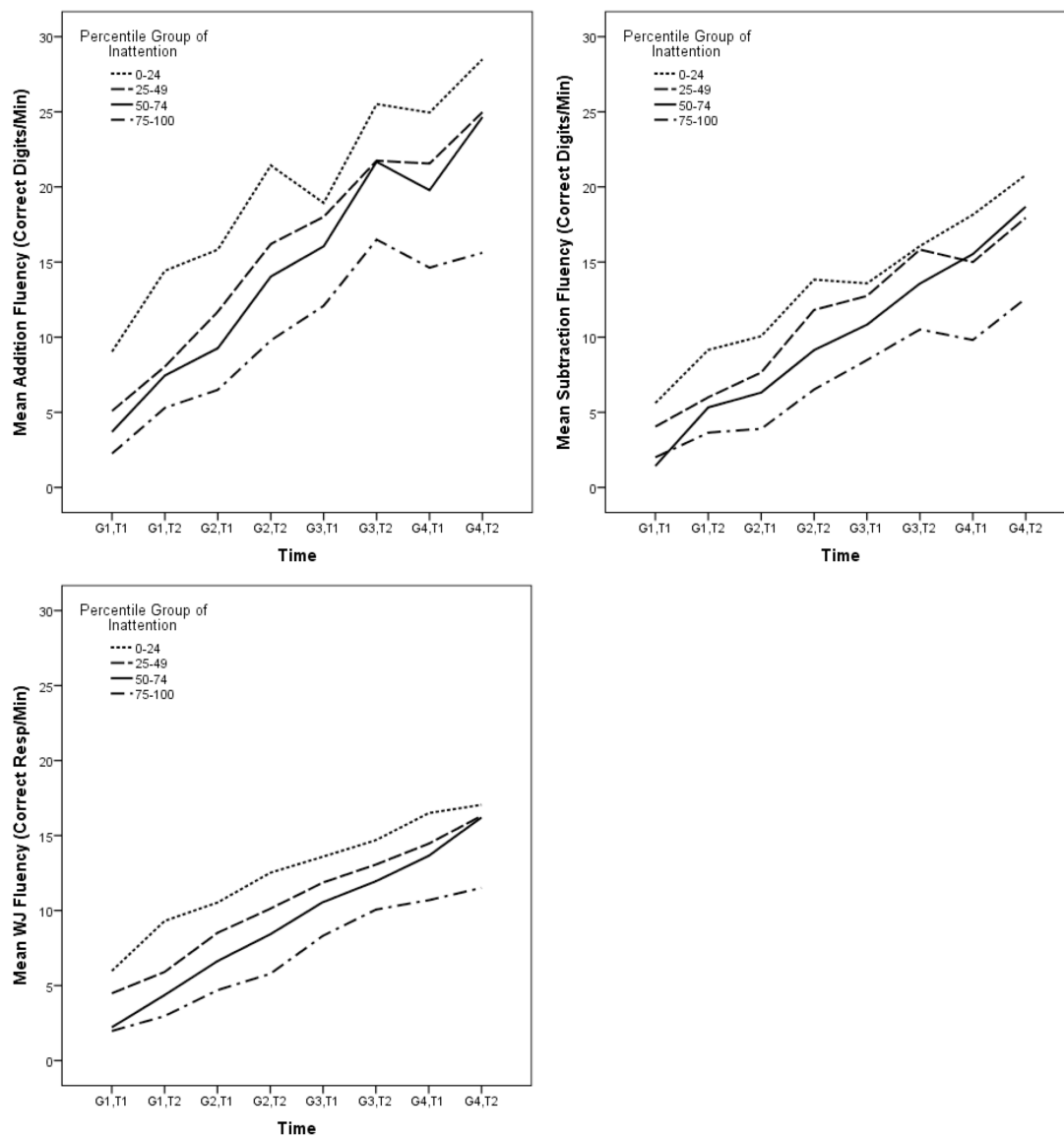
*Figure 6.* Relationship between teacher-rated inattention and math fluency for a) CBM addition,

b) CBM subtraction, and c) WJ-III Math Fluency. Lines depict quartiles of inattention.

# Appendix A

## Scoring guidelines for M-CBM

1. Correct Digits (CD): Each correct digit that a student writes in the correct place value is marked and counted.
2. Full credit for the digits correct is awarded even if work is not shown.
3. For problems that are either incorrect or incomplete, credit is awarded for correct digits that were also in the correct place value.

Examples:

$$\begin{array}{r} 11 \\ +\ \ 7 \\ \hline 18 \end{array}$$
⇧ ⇧

(2 CD possible)

Score = 2 CD

$$\begin{array}{r} 8 \\ +\ \ 7 \\ \hline 14 \end{array}$$
⇧

(2 CD possible)

Score = 1 CD

$$\begin{array}{r} 12 \\ -\ \ 5 \\ \hline 7 \end{array}$$
⇧

(1 CD possible)

Score = 1 CD

$$\begin{array}{r} 12 \\ -\ \ 2 \\ \hline 9 \end{array}$$

(2 CD possible)

Score = 0 CD

# References

Aarnoutse, C., & van Leeuwe, J. (2000). Development of poor and better readers during the elementary school. *Educational Research and Evaluation, 6*(3), 251-278. doi:10.1076/1380-3611(200009)6:3;1-A;FT251

Ackerman, P. L. (1988). Determinants of individual differences during skill acquisition: Cognitive abilities and information processing. *Journal of Experimental Psychology: General, 117*(3), 288.

Ackerman, P. T., Anhalt, J. M., & Dykman, R. A. (1986). Arithmetic automatization failure in children with attention and reading disorders - associations and sequela. *Journal of Learning Disabilities, 19*(4), 222-232.

Ackerman, P. L. (2007). New developments in understanding skilled performance. *Current Directions in Psychological Science, 16*(5), 235-239. doi:10.1111/j.1467-8721.2007.00511.x

Adams, W., & Sheslow, D. (2003). *WRAML2: Wide range assessment of memory and learning: Administration and technical manual* Pearson.

Aitken, M., Martinussen, R., Wolfe, R., & Tannock, R. (2015). Factor structure of the strengths and difficulties questionnaire in a canadian elementary school sample. *Assessment for Effective Intervention, 40*(3), 155-165. doi:10.1177/1534508414560347

Alexander, K. L., Entwisle, D. R., & Olson, L. S. (2001). Schools, achievement, and inequality: A seasonal perspective. *Educational Evaluation and Policy Analysis, 23*(2), 171-191. doi:10.3102/01623737023002171

Alexander, K. L., Entwisle, D. R., & Olson, L. S. (2007a). Lasting consequences of the summer learning gap. *American Sociological Review, 72*(2), 167-180.

Alexander, K. L., Entwisle, D. R., & Olson, L. S. (2007b). Summer learning and its implications: Insights from the beginning school study. *New Directions for Youth Development,* (114), 11-32.

Allinder, R. M., Fuchs, L. S., Fuchs, D., & Hamlett, C. L. (1992). Effects of summer break on math and spelling performance as a function of grade level. *Elementary School Journal, 92*(4), 451-460. doi:10.1086/461701

Allinder, R. M., & Eicher, D. D. (1994). Bouncing back: Regression and recoupment among students with mild disabilities following summer break. *Special Services in the Schools, 8*(2), 129-142.

Alloway, T. P., Elliott, J., & Place, M. (2010). Investigating the relationship between attention and working memory in clinical and community samples. *Child Neuropsychology, 16*(3), 242-254. doi:10.1080/09297040903559655

Alloway, T. P., Gathercole, S. E., Kirkwood, H., & Elliott, J. (2009). The cognitive and behavioral characteristics of children with low working memory. *Child Development, 80*(2), 606-621. doi:10.1111/j.1467-8624.2009.01282.x

Alloway, T. P., Gathercole, S. E., & Pickering, S. J. (2006). Verbal and visuospatial short-term and working memory in children: Are they separable? *Child Development, 77*(6), 1698-1716. doi:10.1111/j.1467-8624.2006.00968.x

Anderman, E. M., Gimbert, B., O'Connell, A. A., & Riegel, L. (2015). Approaches to academic growth assessment. *British Journal of Educational Psychology, 85*(2), 138-153. doi:10.1111/bjep.12053

Anderson, D. (2012). Hierarchical linear modeling (HLM): An introduction to key concepts within cross-sectional and growth modelling frameworks. *Oregon: Behavioral Research & Teaching,*

Ashcraft, M. H., & Fierman, B. A. (1982). Mental addition in third, fourth, and sixth graders. *Journal of Experimental Child Psychology, 33*(2), 216-234. doi:http://dx.doi.org.myaccess.library.utoronto.ca/10.1016/0022-0965(82)90017-0

Ashkenazi, S., Rosenberg-Lee, M., Metcalfe, A. W. S., Swigart, A. G., & Menon, V. (2013). Visuo-spatial working memory is an important source of domain-general vulnerability in the development

of arithmetic cognition. *Neuropsychologia, 51*(11), 2305-2317. doi:10.1016/j.neuropsychologia.2013.06.031

Aunola, K., Leskinen, E., Lerkkanen, M. K., & Nurmi, J. E. (2004). Developmental dynamics of math performance from preschool to grade 2. *Journal of Educational Psychology, 96*(4), 699-713. doi:10.1037/0022-0663.96.4.699

Baddeley, A. (1992). Working memory. *Science, 255*(5044), 556-559. doi:10.1126/science.1736359

Baddeley, A. (1996). The fractionation of working memory. *Proceedings of the National Academy of Sciences of the United States of America, 93*(24), 13468-13472.

Bailey, D. H., Littlefield, A., & Geary, D. C. (2012). The codevelopment of skill at and preference for use of retrieval-based processes for solving addition problems: Individual and sex differences from first to sixth grades. *Journal of Experimental Child Psychology, 113*(1), 78-92. doi:10.1016/j.jecp.2012.04.014

Bailey, D. H., Watts, T. W., Littlefield, A. K., & Geary, D. C. (2014). State and trait effects on individual differences in children's mathematical development. *Psychological Science, 25*(11), 2017-2026. doi:10.1177/0956797614547539

Barnes, M. A., Raghubar, K. P., English, L., Williams, J. M., Taylor, H., & Landry, S. (2014). Longitudinal mediators of achievement in mathematics and reading in typical and atypical development. *Journal of Experimental Child Psychology, 119*, 1-16. doi:10.1016/j.jecp.2013.09.006

Baroody, A. J., Bajwa, N. P., & Eiland, M. (2009). Why can't johnny remember the basic facts? *Developmental Disabilities Research Reviews, 15*(1), 69-79.

Barrouillet, P., & Lepine, R. (2005). Working memory and children's use of retrieval to solve addition problems. *Journal of Experimental Child Psychology, 91*(3), 183-204. doi:10.1016/j.jecp.2005.03.002

Barrouillet, P., Mignon, M., & Thevenot, C. (2008). Strategies in subtraction problem solving in children. *Journal of Experimental Child Psychology, 99*(4), 233-251. doi:10.1016/j.jecp.2007.12.001

Barrouillet, P., & Thevenot, C. (2013). On the problem-size effect in small additions: Can we really discard any counting-based account? *Cognition, 128*(1), 35-44. doi:10.1016/j.cognition.2013.02.018

Bast, J., & Reitsma, P. (1998). Analyzing the development of individual differences in terms of matthew effects in reading: Results from a dutch longitudinal study. *Developmental Psychology, 34*(6), 1373.

Bedard, A., & Tannock, R. (2008). Anxiety, methylphenidate response, and working memory in children with ADHD. *Journal of Attention Disorders, 11*(5), 546-557. doi:10.1177/1087054707311213

Berteletti, I., Man, G., & Booth, J. R. (2015). How number line estimation skills relate to neural activations in single digit subtraction problems. *NeuroImage, 107*, 198-206. doi:10.1016/j.neuroimage.2014.12.011

Berteletti, I., & Booth, J. R. (2015). Perceiving fingers in single-digit arithmetic problems. *Frontiers in Psychology, 6*, 226. doi:10.3389/fpsyg.2015.00226

Bickel, R. (2007). *Multilevel analysis for applied research: It's just regression!*. New York, NY: Guilford Press.

Biederman, J., Mick, E., Fried, R., Aleardi, M., Potter, A., & Herzig, K. (2005). A simulated workplace experience for nonmedicated adults with and without ADHD. *Psychiatric Services, 56*(12), 1617-1620. doi:10.1176/appi.ps.56.12.1617

Binder, C., Haughton, E. & Bateman, B. (2002). Fluency: Achieving true mastery in the learning process. Retrieved from http://special.edschool.virginia.edu/papers/Binder-et-al_Fluency.pdf

Binder, C., Haughton, E., & Van Eyk, D. (1990). Increasing endurance by building fluency: Precision teaching attention span. *Teaching Exceptional Children, 22*(3), 24-27.

Binder, C. (1996). Behavioral fluency: Evolution of a new paradigm. *Behavior Analyst, 19*(2), 163-197.

Binder, C. (2003). Doesn't everybody need fluency? *Performance Improvement, 42*(3), 14-20. doi:10.1002/pfi.4930420304

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

Blackwell, L. S., Trzesniewski, K. H., & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Development, 78*(1), 246-263. doi:10.1111/j.1467-8624.2007.00995.x

Bodovski, K., & Farkas, G. (2007). Mathematics growth in early elementary school: The roles of beginning knowledge, student engagement, and instruction. *Elementary School Journal, 108*(2), 115-130. doi:10.1086/525550

Boissiere, M., Knight, J., & Sabot, R. (1985). Earnings, schooling, ability, and cognitive skills. *American Economic Review, 75*(5), 1016-1030.

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

Burns, M. (2011). Matching math interventions to students' skill deficits: A preliminary investigation of a conceptual and procedural heuristic. *Assessment for Effective Intervention, 36*(4), 210-218. doi:10.1177/1534508411413255

Burns, M. K. (2005). Using incremental rehearsal to increase fluency of single-digit multiplication facts with children identified as learning disabled in mathematics computation. *Education & Treatment of Children, 28*(3), 237-249.

Burns, M. K., Codding, R. S., Boice, C. H., & Lukito, G. (2010). Meta-analysis of acquisition and fluency

    math interventions with instructional and frustration level skills: Evidence for a skill-by-treatment

    interaction. *School Psychology Review, 39*(1), 69-83.

Burns, M. K., Kanive, R., & DeGrande, M. (2012). Effect of a computer-delivered math fact intervention

    as a supplemental intervention for math in third and fourth grades. *Remedial and Special Education,*

    *33*(3), 184-191. doi:10.1177/0741932510381652

Burns, M. K., VanDerHeyden, A. M., & Jiban, C. L. (2006). Assessing the instructional level for

    mathematics: A comparison of methods. *School Psychology Review, 35*(3), 401-418.

Burns, M. K., Walick, C., Simonson, G. R., Dominguez, L., Harelstad, L., Kincaid, A., & Nelson, G. S.

    (2015). Using a conceptual understanding and procedural fluency heuristic to target math

    interventions with students in early elementary. *Learning Disabilities Research & Practice, 30*(2),

    52-60.

Buzhardt, J., Greenwood, C. R., Abbott, M., & Tapia, Y. (2007). Scaling up ClassWide peer tutoring:

    Investigating barriers to wide-scale implementation from a distance. *Learning Disabilities: A*

    *Contemporary Journal, 5*(2), 75-96.

Carr, M., & Davis, H. (2001). Gender differences in arithmetic strategy use: A function of skill and

    preference. *Contemporary Educational Psychology, 26*(3), 330-347. doi:10.1006/ceps.2000.1059

Carr, M., & Jessup, D. L. (1997). Gender differences in first-grade mathematics strategy use: Social and

    metacognitive influences. *Journal of Educational Psychology, 89*(2), 318-328. doi:10.1037/0022-

    0663.89.2.318

Carr, M., & Alexeev, N. (2011). Fluency, accuracy, and gender predict developmental trajectories of

    arithmetic strategies. *Journal of Educational Psychology, 103*(3), 617-631. doi:10.1037/a0023864

Carr, M., Steiner, H. H., Kyser, B., & Biddlecomb, B. (2008). A comparison of predictors of early

emerging gender differences in mathematics competency. *Learning and Individual Differences,*

*18*(1), 61-75. doi:10.1016/j.lindif.2007.04.005

Carr, M., Taasoobshirazi, G., Stroud, R., & Royer, J. M. (2011). Combined fluency and cognitive

strategies instruction improves mathematics achievement in early elementary school. *Contemporary*

*Educational Psychology, 36*(4), 323-333. doi:10.1016/j.cedpsych.2011.04.002

Cates, G. L. (2005). Effects of peer versus computer-assisted drill on mathematics response rates.

*Psychology in the Schools, 42*(6), 637-646.

Cates, G. L., & Rhymer, K. N. (2003). Examining the relationship between mathematics anxiety and

mathematics performance: An instructional hierarchy perspective. *Journal of Behavioral Education,*

*12*(1), 23-34.

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal

recall tasks: A review and quantitative synthesis. *Psychological Bulletin, 132*(3), 354.

Cerda, C. A., Im, M. H., & Hughes, J. N. (2014). Learning-related skills and academic achievement in

academically at-risk first graders. *Journal of Applied Developmental Psychology, 35*(5), 433-443.

doi:10.1016/j.appdev.2014.08.001

Cheung, A. C. K., & Slavin, R. E. (2013). The effectiveness of educational technology applications for

enhancing mathematics achievement in K-12 classrooms: A meta-analysis. *Educational Research*

*Review, 9*, 88-113. doi:10.1016/j.edurev.2013.01.001

Chong, S. L., & Siegel, L. S. (2008). Stability of computational deficits in math learning disability from

second through fifth grades. *Developmental Neuropsychology, 33*(3), 300-317.

doi:10.1080/87565640801982387

Christ, T. J., Van Norman, E. R., & Nelson, P. M. (2016). Foundations of fluency-based assessments in

behavioral and psychometric paradigms. In K. D. Cummings, & Y. M. Petscher (Eds.), *The fluency

construct: Curriculum-based measurement concepts and applications* (pp. 143-163). New York:

Springer.

Christ, T. J., & Vining, O. (2006). Curriculum-based measurement procedures to develop multiple-skill

mathematics computation probes: Evaluation of random and stratified stimulus-set arrangements.

*School Psychology Review, 35*(3), 387-400.

Christ, T., Scullin, S., Tolbize, A., & Jiban, C. (2008). Implications of recent research. *Assessment for

Effective Intervention, 33*(4), 198-205. doi:10.1177/1534508407313480

Clarke, B., Nelson, N., & Shanley, L. (2016). Chapter 3: Mathematics fluency - more than the weekly

timed test. In K. D. Cummings, & Y. M. Petscher (Eds.), *The fluency construct: Curriculum-based

measurement concepts and applications* (pp. 67-89). New York, NY: Springer.

Clarke, B., Doabler, C. T., Nelson, N. J., & Shanley, C. (2015). Effective instructional strategies for

kindergarten and first-grade students at risk in mathematics. *Intervention in School and Clinic, 50*(5),

257-265. doi:10.1177/1053451214560888

Codding, R. S., Shiyko, M., Russo, M., Birch, S., Fanning, E., & Jaspen, D. (2007). Comparing

mathematics interventions: Does initial level of fluency predict intervention effectiveness? *Journal

of School Psychology, 45*(6), 603-617. doi:10.1016/j.jsp.2007.06.005

Codding, R. S., Burns, M. K., & Lukito, G. (2011). Meta-analysis of mathematic basic-fact fluency

interventions: A component analysis. *Learning Disabilities Research & Practice, 26*(1), 36-47.

doi:10.1111/j.1540-5826.2010.00323.x

Codding, R. S., Hilt-Panahon, A., Panahon, C. J., & Benson, J. L. (2009, 05; 2016/9). Addressing

    mathematics computation problems: A review of simple and moderate intensity interventions.*32*,

    279+.

Collins, L. M. (2006). Analysis of longitudinal data: The integration of theoretical model, temporal

    design, and statistical model. *Annual Review of Psychology, 57*, 505-528.

    doi:10.1146/annurev.psych.57.102904.190146

Colom, R., Shih, P. C., Flores-Mendoza, C., & Quiroga, M. A. (2006). The real relationship between

    short-term memory and working memory. *Memory, 14*(7), 804-813.

    doi:10.1080/09658210600680020

Cooper, H., Nye, B., Charlton, K., Lindsay, J., & Greathouse, S. (1996). The effects of summer vacation

    on achievement test scores: A narrative and meta-analytic review. *Review of Educational Research,*

    *66*(3), 227-268. doi:10.3102/00346543066003227

Cooper, G., & Sweller, J. (1987). Effects of schema acquisition and rule automation on mathematical

    problem-solving transfer. *Journal of Educational Psychology, 79*(4), 347.

Cooper, H., Charlton, K., Valentine, J. C., & Muhlenbruck, L. (2000). Making the most of summer

    school: A meta-analytic and narrative review. *Monographs of the Society for Research in Child*

    *Development, 65*(1), i-v, 1-118; discussion 119-27.

Cooper, H., Valentine, J. C., Charlton, K., & Melson, A. (2003). The effects of modified school calendars

    on student achievement and on school and community attitudes. *Review of Educational Research,*

    *73*(1), 1-52. doi:10.3102/00346543073001001

Cowan, R., Donlan, C., Shepherd, D., Cole-Fletcher, R., Saxton, M., & Hurry, J. (2011). Basic calculation

    proficiency and mathematics achievement in elementary school children. *Journal of Educational*

    *Psychology, 103*(4), 786-803. doi:10.1037/a0024556

Currie, J., & Stabile, M. (2006). Child mental health and human capital accumulation: The case of ADHD. *Journal of Health Economics, 25*(6), 1094-1118. doi:10.1016/j.jhealeco.2006.03.001

Daly, E. J. I., Martens, B. K., Barnett, D., Witt, J. C., & Olson, S. C. (2007). Varying intervention delivery in response to intervention: Confronting and resolving challenges with measurement, instruction, and intensity. *School Psychology Review, 36*(4), 562-581.

Davies, S., & Aurini, J. (2014). Summer learning program 2013 quantitative and qualitative research results. Retrieved from http://www.ontariodirectors.ca/Summer_Learning-2014/40821_SummerLearning_Report_E.pdf

Davies, S., & Aurini, J. (2010). The ontario summer literacy learning project. Retrieved from http://www.edu.gov.on.ca/eng/literacynumeracy/research/summerliteracy.pdf

Davies, S., & Aurini, J. (2013). Summer learning inequality in ontario. *Canadian Public Policy-Analyse De Politiques, 39*(2), 287-307.

Davies, S., Aurini, J., Milne, E., & Jean-Pierre, J. (2015). Les effets des programmes d'été de littératie: Les théories sur les opportunités d'apprentissage et les élèves « non traditionnels » dans les écoles de langue française en ontario. *Canadian Journal of Sociology (Online), 40*(2), 189-222.

Davis-Kean, P. E. (2005). The influence of parent education and family income on child achievement: The indirect role of parental expectations and the home environment. *Journal of Family Psychology, 19*(2), 294-304. doi:10.1037/0893-3200.19.2.294

De Smedt, B., Torbeyns, J., Stassens, N., Ghesquiere, P., & Verschaffel, L. (2010). Frequency, efficiency and flexibility of indirect addition in two learning environments. *Learning and Instruction, 20*(3), 205-215.

De Smedt, B., Holloway, I. D., & Ansari, D. (2011). Effects of problem size and arithmetic operation on brain activation during calculation in children with varying levels of arithmetical fluency. *NeuroImage, 57*(3), 771-781. doi:10.1016/j.neuroimage.2010.12.037

De Smedt, B., Reynvoet, B., Swillen, A., Verschaffel, L., Boets, B., & Ghesquiere, P. (2009). Basic number processing and difficulties in single-digit arithmetic: Evidence from velo-cardio-facial syndrome. *Cortex, 45*(2), 177-188. doi:10.1016/j.cortex.2007.06.003

De Smedt, B., Taylor, J., Archibald, L., & Ansari, D. (2010). How is phonological processing related to individual differences in children's arithmetic skills? *Developmental Science, 13*(3), 508-520. doi:10.1111/j.1467-7687.2009.00897.x

de Winter, J. C. (2013). Using the student's t-test with extremely small sample sizes. *Practical Assessment, Research & Evaluation, 18*(10), 1-12.

Dehaene, S. (1992). Varieties of numerical abilities. *Cognition, 44*(1-2), 1-42. doi:10.1016/0010-0277(92)90049-N

Dehaene, S., & Cohen, L. (1997). Cerebral pathways for calculation: Double dissociation between rote verbal and quantitative knowledge of arithmetic. *Cortex, 33*(2), 219-250. doi:10.1016/S0010-9452(08)70002-9

Dehaene, S., Piazza, M., Pinel, P., & Cohen, L. (2003). Three parietal circuits for number processing. *Cognitive Neuropsychology, 20*(3-6), 487-506. doi:10.1080/02643290244000239

Deno, S. L. (2003). Developments in curriculum-based measurement. *The Journal of Special Education, 37*(3), 184-192.

Deno, S. L. (1985). Curriculum-based measurement - the emerging alternative. *Exceptional Children, 52*(3), 219-232.

Diamantopoulou, S., Rydell, A., Thorell, L. B., & Bohlin, G. (2007). Impact of executive functioning and

    symptoms of attention deficit hyperactivity disorder on children's peer relations and school

    performance. *Developmental Neuropsychology, 32*(1), 521-542.

Doggett, R., Henington, C., & Johnson-Gros, K. (2006). Math to mastery: A direct instruction remedial

    math intervention designed to increase student fluency with basic math facts. *Unpublished*

    *Manuscript, Mississippi State University,*

Downey, D. B., von Hippel, P. T., & Broh, B. A. (2004). Are schools the great equalizer? cognitive

    inequality during the summer months and the school year. *American Sociological Review, 69*(5),

    613-635.

Duckworth, A. L., & Seligman, M. E. P. (2006). Self-discipline gives girls the edge: Gender in self-

    discipline, grades, and achievement test scores. *Journal of Educational Psychology, 98*(1), 198-208.

    doi:10.1037/0022-0663.98.1.198

Duhon, G. J., House, S., Hastings, K., Poncy, B., & Solomon, B. (2015). Adding immediate feedback to

    explicit timing: An option for enhancing treatment intensity to improve mathematics fluency.

    *Journal of Behavioral Education, 24*(1), 74-87. doi:10.1007/s10864-014-9203-y

Duncan, G. J., Claessens, A., Huston, A. C., Pagani, L. S., Engel, M., Sexton, H., . . . Duckworth, K.

    (2007). School readiness and later achievement. *Developmental Psychology, 43*(6), 1428-1446.

    doi:10.1037/0012-1649.43.6.1428

DuPaul, G. J., & Eckert, T. L. (1998). Academic interventions for students with attention-

    deficit/hyperactivity disorder: A review of the literature. *Reading & Writing Quarterly, 14*(1), 59-82.

    doi:10.1080/1057356980140104

DuPaul, G., Ervin, R., Hook, C., & McGoey, K. (1998). Peer tutoring for children with attention deficit

    hyperactivity disorder: Effects on classroom behavior and academic performance. *Journal of Applied*

    *Behavior Analysis, 31*(4), 579-592. doi:10.1901/jaba.1998.31-579

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term

    memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental*

    *Psychology-General, 128*(3), 309-331. doi:10.1037/0096-3445.128.3.309

Entwisle, D. R., & Alexander, K. L. (1990). Beginning school math competence - minority and majority

    comparisons. *Child Development, 61*(2), 454-471. doi:10.1111/j.1467-8624.1990.tb02792.x

Entwisle, D. R., & Alexander, K. L. (1992). Summer setback - race, poverty, school composition, and

    mathematics achievement in the 1st 2 years of school. *American Sociological Review, 57*(1), 72-84.

    doi:10.2307/2096145

Espin, C., Deno, S., Maruyama, G., & Cohen, C. (1989). The basic academic skills sample (BASS): An

    instrument for the screening and identification of children at risk for failure in regular education

    classrooms. *Annual Meeting of American Educational Research Association, San Francisco, CA,*

Fabiano, G. A., Schatz, N. K., & Pelham, W. E., Jr. (2014). Summer treatment programs for youth with

    ADHD. *Child and Adolescent Psychiatric Clinics of North America, 23*(4), 757-+.

    doi:10.1016/j.chc.2014.05.012

Fayol, M., & Thevenot, C. (2012). The use of procedural knowledge in simple addition and subtraction

    problems. *Cognition, 123*(3), 392-403. doi:10.1016/j.cognition.2012.02.008

Fitzpatrick, C., & Pagani, L. S. (2013). Task-oriented kindergarten behavior pays off in later childhood.

    *Journal of Developmental and Behavioral Pediatrics, 34*(2), 94-101.

Foegen, A., Jiban, C., & Deno, S. (2007). Progress monitoring measures in mathematics. *The Journal of*

    *Special Education, 41*(2), 121-139. doi:10.1177/00224669070410020101

Friso-van den Bos, I., van der Ven, S. H. G., Kroesbergen, E. H., & van Luit, J. E. H. (2013). Working

    memory and mathematics in primary school children: A meta-analysis. *Educational Research*

    *Review, 10*, 29-44. doi:10.1016/j.edurev.2013.05.003

Fuchs, L. S., & Deno, S. L. (1991). Paradigmatic distinctions between instructionally relevant

    measurement models. *Exceptional Children, 57*(6), 488-500.

Fuchs, L. S., Compton, D. L., Fuchs, D., Paulsen, K., Bryant, J. D., & Hamlett, C. L. (2005). The

    prevention, identification, and cognitive determinants of math difficulty. *Journal of Educational*

    *Psychology, 97*(3), 493-513. doi:10.1037/0022-0663.97.3.493

Fuchs, L. S., Fuchs, D., Compton, D. L., Powell, S. R., Seethaler, P. M., Capizzi, A. M., . . . Fletcher, J.

    M. (2006). The cognitive correlates of third-grade skill in arithmetic, algorithmic computation, and

    arithmetic word problems. *Journal of Educational Psychology, 98*(1), 29-43. doi:10.1037/0022-

    0663.98.1.29

Fuchs, L. S., Fuchs, D., Hamlett, C. L., Walz, L., & Germann, G. (1993). Formative evaluation of

    academic progress - how much growth can we expect. *School Psychology Review, 22*(1), 27-48.

Fuchs, L. S., Fuchs, D., Stuebing, K., Fletcher, J. M., Hamlett, C. L., & Lambert, W. (2008). Problem

    solving and computational skill: Are they shared or distinct aspects of mathematical cognition?

    *Journal of Educational Psychology, 100*(1), 30-47. doi:10.1037/0022-0663.100.1.30

Fuchs, L. S., Geary, D. C., Compton, D. L., Fuchs, D., Hamlett, C. L., & Bryant, J. D. (2010). The

    contributions of numerosity and domain-general abilities to school readiness. *Child Development,*

    *81*(5), 1520-1533. doi:10.1111/j.1467-8624.2010.01489.x

Fuchs, L. S., Powell, S. R., Seethaler, P. M., Fuchs, D., Hamlett, C. L., Cirino, P. T., & Fletcher, J. M.

    (2010). A framework for remediating number combination deficits. *Exceptional Children, 76*(2),

    135-156.

Fuhs, M. W., Hornburg, C. B., & McNeil, N. M. (2016). Specific early number skills mediate the

   association between executive functioning skills and mathematics achievement. *Developmental*

   *Psychology, 52*(8), 1217-1235.

   doi:http://dx.doi.org.myaccess.library.utoronto.ca/10.1037/dev0000145

Garner, A. A., O'Connor, B. C., Narad, M. E., Tamm, L., Simon, J., & Epstein, J. N. (2013). The

   relationship between ADHD symptom dimensions, clinical correlates, and functional impairments.

   *Journal of Developmental and Behavioral Pediatrics, 34*(7), 469-477.

Gau, S. S. (2011). Childhood trajectories of inattention symptoms predicting educational attainment in

   young adults. *American Journal of Psychiatry, 168*(11), 1131-1133.

Geary, D. (2004). Mathematics and learning disabilities. *Journal of Learning Disabilities, 37*(1), 4-15.

   doi:10.1177/00222194040370010201

Geary, D. C. (1993). Mathematical disabilities - cognitive, neuropsychological, and genetic components.

   *Psychological Bulletin, 114*(2), 345-362. doi:10.1037//0033-2909.114.2.345

Geary, D. C., BowThomas, C. C., Liu, F., & Siegler, R. S. (1996). Development of arithmetical

   competencies in chinese and american children: Influence of age, language, and schooling. *Child*

   *Development, 67*(5), 2022-2044. doi:10.1111/j.1467-8624.1996.tb01841.x

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*(5), 787-797. doi:10.1037/0012-1649.27.5.787

Geary, D. C., & Burlinghamdubree, M. (1989). External validation of the strategy choice model for

   addition. *Journal of Experimental Child Psychology, 47*(2), 175-192. doi:10.1016/0022-

   0965(89)90028-3

Geary, D. C., Hoard, M. K., Byrd-Craven, J., & DeSoto, M. C. (2004). Strategy choices in simple and

    complex addition: Contributions of working memory and counting knowledge for children with

    mathematical disability. *Journal of Experimental Child Psychology, 88*(2), 121-151.

    doi:10.1016/j.jecp.2004.03.002

Geary, D. C., Saults, S. J., Liu, F., & Hoard, M. K. (2000). Sex differences in spatial cognition,

    computational fluency, and arithmetical reasoning. *Journal of Experimental Child Psychology,*

    *77*(4), 337-353. doi:10.1006/jecp.2000.2594

Geary, D. C. (2011a). Cognitive predictors of achievement growth in mathematics: A 5-year longitudinal

    study. *Developmental Psychology, 47*(6), 1539-1552. doi:10.1037/a0025510

Geary, D. C. (2011b). Consequences, characteristics, and causes of mathematical learning disabilities and

    persistent low achievement in mathematics. *Journal of Developmental and Behavioral Pediatrics,*

    *32*(3), 250-263. doi:10.1097/DBP.0b013e318209edef

Geary, D. C., Hoard, M. K., & Bailey, D. H. (2012). Fact retrieval deficits in low achieving children and

    children with mathematical learning disability. *Journal of Learning Disabilities, 45*(4), 291-307.

    doi:10.1177/0022219410392046

Geary, D. C., Hoard, M. K., Byrd-Craven, J., Nugent, L., & Numtee, C. (2007). Cognitive mechanisms

    underlying achievement deficits in children with mathematical learning disability. *Child*

    *Development, 78*(4), 1343-1359. doi:10.1111/j.1467-8624.2007.01069.x

Geary, D. C., Hoard, M. K., & Nugent, L. (2012). Independent contributions of the central executive,

    intelligence, and in-class attentive behavior to developmental change in the strategies used to solve

    addition problems. *Journal of Experimental Child Psychology, 113*(1), 49-65.

    doi:10.1016/j.jecp.2012.03.003

Geary, D. C., Hoard, M. K., Nugent, L., & Bailey, D. H. (2012). Mathematical cognition deficits in
children with learning disabilities and persistent low achievement: A five-year prospective study.
*Journal of Educational Psychology, 104*(1), 206-223. doi:10.1037/a0025398

Gershon, J. (2002). A meta-analytic review of gender differences in ADHD. *Journal of Attention
Disorders, 5*(3), 143-54. doi:10.1177/108705470200500302

Gersten, R., & Chard, D. (1999). Number sense. *The Journal of Special Education, 33*(1), 18-28.
doi:10.1177/002246699903300102

Giannopulu, I., Escolano, S., Cusin, F., Citeau, H., & Dellatolas, G. (2008). Teachers' reporting of
behavioural problems and cognitive-academic performances in children aged 5-7 years. *British
Journal of Educational Psychology, 78*, 127-147. doi:10.1348/000709907X204372

Gickling, E. E., & Thompson, V. P. (1985). A personal view of curriculum-based assessment.
*Exceptional Children, 52*(3), 205-218.

Golbeck, A. L., Ahlers-Schmidt, C. R., Paschal, A. M., & Dismuke, S. E. (2005). A definition and
operational framework for health numeracy. *American Journal of Preventive Medicine, 29*(4), 375-
376. doi:10.1016/j.amepre.2005.06.012

Gold, A. B., Ewing-Cobbs, L., Cirino, P., Fuchs, L. S., Stuebing, K. K., & Fletcher, J. M. (2013).
Cognitive and behavioral attention in children with math difficulties. *Child Neuropsychology, 19*(4),
420-437. doi:10.1080/09297049.2012.690371

Gouvernement du Québec, Ministère de l'Éducation. (2001). Québec education program, approved
version. Retrieved from
http://www1.education.gouv.qc.ca/sections/programmeFormation/primaire/pdf/educprg2001/educpr
g2001.pdf

Graney, S. B., Missall, K. N., Martinez, R. S., & Bergstrom, M. (2009). A preliminary investigation of within-year growth patterns in reading and mathematics curriculum-based measures. *Journal of School Psychology, 47*(2), 121-142. doi:10.1016/j.jsp.2008.12.001

Gray, S. A., Rogers, M., Martinussen, R., & Tannock, R. (2015). Longitudinal relations among inattention, working memory, and academic achievement: Testing mediation and the moderating role of gender. *Peerj, 3*, e939. doi:10.7717/peerj.939

Gunderson, E. A., Ramirez, G., Beilock, S. L., & Levine, S. C. (2012). The relation between spatial skill and early number knowledge: The role of the linear number line. *Developmental Psychology, 48*(5), 1229-1241. doi:10.1037/a0027433

Han, W. (2008). The academic trajectories of children of immigrants and their school environments. *Developmental Psychology, 44*(6), 1572-1590. doi:10.1037/a0013886

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*(3), 615-626. doi:10.1037/0022-0663.93.3.615

Haring, N. G., & Eaton, M. D. (1978). Systematic instructional procedures: An instructional hierarchy. In N. G. Haring, T. C. Lovitt, M. D. Eaton & C. L. Hansen (Eds.), *The fourth R: Research in the classroom.* (pp. 23-40). Columbus, OH: Merrill.

Haring, N. G., Lovitt, T. C., Eaton, M. D., & Hansen, C. L. (1978). The 4th R: Research in the classroom. *Columbus, OH: Charles E.Merrill,*

Harrison, J. R., Bunford, N., Evans, S. W., & Owens, J. S. (2013). Educational accommodations for students with behavioral challenges: A systematic review of the literature. *Review of Educational Research, 83*(4), 551-597. doi:10.3102/0034654313497517

Hart, S. A., Petrill, S. A., & Thompson, L. A. (2010). A factorial analysis of timed and untimed measures of mathematics and reading abilities in school aged twins. *Learning and Individual Differences, 20*(2), 63-69. doi:10.1016/j.lindif.2009.10.004

Hart, S. A., Petrill, S. A., Thompson, L. A., & Plomin, R. (2009). The ABCs of math: A genetic analysis of mathematics and its links with reading ability and general cognitive ability. *Journal of Educational Psychology, 101*(2), 388-402. doi:10.1037/a0015115

Hartnedy, S. L., Mozzoni, M. P., & Fahoum, Y. (2005). The effect of fluency training on math and reading skills in neuropsychiatric diagnosis children: A multiple baseline. *Behavioral Interventions, 20*(1), 27-36. doi:10.1002/bin.167

Hawkins, R. O., Collins, T., Hernan, C., & Flowers, E. (2016). Using computer-assisted instruction to build math fact fluency an implementation guide. *Intervention in School and Clinic, ,* 1053451216644827.

Hecht, S. A., Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (2001). The relations between phonological processing abilities and emerging individual differences in mathematical computation skills: A longitudinal study from second to fifth grades. *Journal of Experimental Child Psychology, 79*(2), 192-227. doi:10.1006/jecp.2000.2586

Henry, V. J., & Brown, R. S. (2008). First-grade basic facts: An investigation into teaching and learning of an accelerated, high-demand memorization standard. *Journal for Research in Mathematics Education, 39*(2), 153-183. doi:10.2307/30034895

Heyns, B. (1987). Schooling and cognitive-development - is there a season for learning. *Child Development, 58*(5), 1151-1160.

Hintze, J. M., Christ, T. J., & Keller, L. A. (2002). The generalizability of CBM survey-level mathematics assessments: Just how many samples do we need? *School Psychology Review, 31*(4), 514-528.

Holmberg, K., & Bolte, S. (2014). Do symptoms of ADHD at ages 7 and 10 predict academic outcome at age 16 in the general population? *Journal of Attention Disorders, 18*(8), 635-645. doi:10.1177/1087054712452136

Holmes, J., & Adams, J. W. (2006). Working memory and children's mathematical skills: Implications for mathematical development and mathematics curricula. *Educational Psychology, 26*(3), 339-366. doi:10.1080/01443410500341056

Hooper, S. R., Roberts, J., Sideris, J., Burchinal, M., & Zeisel, S. (2010). Longitudinal predictors of reading and math trajectories through middle school for african american versus caucasian students across two samples. *Developmental Psychology, 46*(5), 1018.

Hosp, M., Hosp, J., & Howell, K. (2007). The ABCs of CBM. *A Practical Guide to Curriculum-Based Measurement,*

Huang, F. L., Moon, T. R., & Boren, R. (2014). Are the reading rich getting richer? testing for the presence of the matthew effect. *Reading & Writing Quarterly, 30*(2), 95-115. doi:10.1080/10573569.2013.789784

Hubber, P. J., Gilmore, C., & Cragg, L. (2014). The roles of the central executive and visuospatial storage in mental arithmetic: A comparison across strategies. *Quarterly Journal of Experimental Psychology, 67*(5), 936-954. doi:10.1080/17470218.2013.838590

Huttenlocher, J., Jordan, N. C., & Levine, S. C. (1994). A mental model for early arithmetic. *Journal of Experimental Psychology-General, 123*(3), 284-296. doi:10.1037/0096-3445.123.3.284

IMB Corp. (2010). *IBM SPSS statistics for windows, version 19.0*. Armonk, NY: IBM Corp.

Imbo, I., & Vandierendonck, A. (2008). Effects of problem size, operation, and working-memory span on simple-arithmetic strategies: Differences between children and adults? *Psychological Research, 72*(3), 331-346.

Imbo, I., & Vandierendonck, A. (2007). The development of strategy use in elementary school children: Working memory and individual differences. *Journal of Experimental Child Psychology, 96*(4), 284-309. doi:10.1016/j.jecp.2006.09.001

Imbo, I., Vandierendonck, A., & Rosseel, Y. (2007). The influence of problem features and individual differences on strategic performance in simple arithmetic. *Memory & Cognition, 35*(3), 454-463. doi:10.3758/BF03193285

Jordan, J., Mulhern, G., & Wylie, J. (2009). Individual differences in trajectories of arithmetical development in typically achieving 5-to 7-year-olds. *Journal of Experimental Child Psychology, 103*(4), 455-468. doi:10.1016/j.jecp.2009.01.011 ER

Jordan, N. C., Hanich, L. B., & Kaplan, D. (2003). Arithmetic fact mastery in young children: A longitudinal investigation. *Journal of Experimental Child Psychology, 85*(2), 103-119. doi:10.1016/S0022-0965(03)00032-8

Jordan, N. C., Kaplan, D., & Hanich, L. B. (2002). Achievement growth in children with learning difficulties in mathematics: Findings of a two-year longitudinal study. *Journal of Educational Psychology, 94*(3), 586-597. doi:10.1037//0022-1663.3.586

Jordan, N. C., Kaplan, D., Olah, L. N., & Locuniak, M. N. (2006). Number sense growth in kindergarten: A longitudinal investigation of children at risk for mathematics difficulties. *Child Development, 77*(1), 153-175. doi:10.1111/j.1467-8624.2006.00862.x

Jordan, N. C., Kaplan, D., Ramineni, C., & Locuniak, M. N. (2009). Early math matters: Kindergarten number competence and later mathematics outcomes. *Developmental Psychology, 45*(3), 850-867. doi:10.1037/a0014939

Jordan, N., & Montani, T. (1997). Cognitive arithmetic and problem solving. *Journal of Learning Disabilities, 30*(6), 624-634. doi:10.1177/002221949703000606

Jordan, N., Hanich, L., & Kaplan, D. (2003). A longitudinal study of mathematical competencies in

    children with specific mathematics difficulties versus children with comorbid mathematics and

    reading difficulties. *Child Development, 74*(3), 834-850. doi:10.1111/1467-8624.00571

Joseph, L. M., Konrad, M., Cates, G., Vajcner, T., Eveleigh, E., & Fishley, K. M. (2012). A meta-analytic

    review of the cover-copy-compare and variations of this self-management procedure. *Psychology in*

    *the Schools, 49*(2), 122-136. doi:10.1002/pits.20622

Judge, S., & Watson, S. M. R. (2011). Longitudinal outcomes for mathematics achievement for students

    with learning disabilities. *Journal of Educational Research, 104*(3), 147-157.

    doi:10.1080/00220671003636729

Kamii, C., & Lewis, B. A. (2003). Single-digit subtraction with fluency. *Teaching Children Mathematics,*

    *10*(4), 230-236. doi:12/2003

Kamii, C., Lewis, B. A., & Kirkland, L. D. (2001). Fluency in subtraction compared with addition.

    *Journal of Mathematical Behavior, 20*(1), 33-42. doi:10.1016/S0732-3123(01)00060-8

Kavale, K. A., & Forness, S. R. (2000). Policy decisions in special education: The role of meta-analysis.

    In E. Gersten, P. Schiller & S. Vaughn (Eds.), *Contemporary special education research: Syntheses*

    *of knowledge base on critical instructional issues*. (pp. 281-326). Mahwah, NJ: Erlbaum.

Keller-Margulis, M., Mercer, S., & Shapiro, E. (2014). Differences in growth on math curriculum-based

    measures using triannual benchmarks. *Assessment for Effective Intervention, 39*(3), 146-155.

    doi:10.1177/1534508412452750

Kochanska, G., Murray, K. T., & Harlan, E. T. (2000). Effortful control in early childhood: Continuity

    and change, antecedents, and implications for social development. *Developmental Psychology,*

    *36*(2), 220.

Kohli, N., Harring, J. R., & Hancock, G. R. (2013). Piecewise linear-linear latent growth mixture models with unknown knots. *Educational and Psychological Measurement, 73*(6), 935-955. doi:10.1177/0013164413496812

Kohli, N., Sullivan, A. L., Sadeh, S., & Zopluoglu, C. (2015). Longitudinal mathematics development of students with learning disabilities and students without disabilities: A comparison of linear, quadratic, and piecewise linear mixed effects models. *Journal of School Psychology, 53*(2), 105-120. doi:10.1016/j.jsp.2014.12.002

Krinzinger, H., Wood, G., & Willmes, K. (2012). What accounts for individual and gender differences in the multi-digit number processing of primary school children? *Zeitschrift Für Psychologie/Journal of Psychology, 220*(2), 78-89. doi:http://dx.doi.org.myaccess.library.utoronto.ca/10.1027/2151-2604/a000099

Kyttala, M., Kanerva, K., & Kroesbergen, E. (2015). Training counting skills and working memory in preschool. *Scandinavian Journal of Psychology, 56*(4), 363-370. doi:10.1111/sjop.12221

Lachance, J. A., & Mazzocco, M. M. M. (2006). A longitudinal analysis of sex differences in math and spatial skills in primary school age children. *Learning and Individual Differences, 16*(3), 195-216. doi:10.1016/j.lindif.2005.12.001

Laski, E. V., Casey, B. M., Yu, Q., Dulaney, A., Heyman, M., & Dearing, E. (2013). Spatial skills as a predictor of first grade girls' use of higher level arithmetic strategies. *Learning and Individual Differences, 23*, 123-130. doi:10.1016/j.lindif.2012.08.001

Leahey, E., & Guo, G. (2001). Gender differences in mathematical trajectories. *Social Forces, 80*(2), 713-732. doi:10.1353/sof.2001.0102

Learning Disabilities Association of Ontario. (2001). *Learning disabilities: A new definition.* Retrieved from http://www.ldao.ca/documents/Definition_and_Suporting%20Document_2001.pdf

Lee, J. (2010). Tripartite growth trajectories of reading and math achievement tracking national academic progress at primary, middle, and high school levels. *American Educational Research Journal, 47*(4), 800-832.

Lee, S. S., & Hinshaw, S. P. (2006). Predictors of adolescent functioning in girls with attention deficit hyperactivity disorder (ADHD): The role of childhood ADHD, conduct problems, and peer status. *Journal of Clinical Child and Adolescent Psychology, 35*(3), 356-368.

LeFevre, J., DeStefano, D., Penner-Wilger, M., & Daley, K. E. (2006). Selection of procedures in mental subtraction. *Canadian Journal of Experimental Psychology/Revue Canadienne De Psychologie Expérimentale, 60*(3), 209.

LeFevre, J., Penner-Wilger, M., Pyke, A. A., Shanahan, T., & Deslauriers, W. A. (2014). Putting two and two together: Declines in arithmetic fluency among young canadian adults, 1993 to 2005. *Ottawa, ON: Carleton University,*

LeFevre, J., Berrigan, L., Vendetti, C., Kamawar, D., Bisanz, J., Skwarchuk, S., & Smith-Chant, B. L. (2013). The role of executive attention in the acquisition of mathematical skills for children in grades 2 through 4. *Journal of Experimental Child Psychology, 114*(2), 243-261. doi:10.1016/j.jecp.2012.10.005

LeFevre, J., Fast, L., Skwarchuk, S., Smith-Chant, B. L., Bisanz, J., Kamawar, D., & Penner-Wilger, M. (2010). Pathways to mathematics: Longitudinal predictors of performance. *Child Development, 81*(6), 1753-1767. doi:10.1111/j.1467-8624.2010.01508.x

LeFevre, J., Jiménez Lira, C., Sowinski, C., Cankaya, O., Kamawar, D., & Skwarchuk, S. (2013). Charting the role of the number line in mathematical development. *Frontiers in Psychology, 4*, 641. doi:10.3389/fpsyg.2013.00641

LeFevre, J., Sadesky, G. S., Bisanz, J., LeFevre, J., Sadesky, G. S., & Bisanz, J. (1996). Selection of procedures in mental addition: Reassessing the problem size effect in adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*(1), 216-230. doi:10.1037/0278-7393.22.1.216

Lemaire, P., & Siegler, R. S. (1995). Four aspects of strategic change: Contributions to children's learning of multiplication. *Journal of Experimental Psychology: General, 124*(1), 83.

Lewandowski, L. J., Lovett, B. J., Parolin, R., Gordon, M., & Codding, R. S. (2007). Extended time accommodations and the mathematics performance of students with and without ADHD. *Journal of Psychoeducational Assessment, 25*(1), 17-28. doi:10.1177/0734282906291961

Li, Y., & Geary, D. C. (2013). Developmental gains in visuospatial memory predict gains in mathematics achievement. *Plos One, 8*(7), e70160. doi:10.1371/journal.pone.0070160

Li-Grining, C. P., Votruba-Drzal, E., Maldonado-Carreno, C., & Haas, K. (2010). Children's early approaches to learning and academic trajectories through fifth grade. *Developmental Psychology, 46*(5), 1062-1077. doi:10.1037/a0020066

Lindberg, S. M., Hyde, J. S., Petersen, J. L., & Linn, M. C. (2010). New trends in gender and mathematics performance: A meta-analysis. *Psychological Bulletin, 136*(6), 1123-1135. doi:10.1037/a0021276

Lindsley, O. R. (1996). Is fluency free-operant response-response chaining? *Behavior Analyst, 19*(2), 211-224.

Little, T. D., Card, N. A., Preacher, K. J., & McConnell, E. (2009). Modeling longitudinal data from research on adolescence. In R. Lerner, & L. Steinberg (Eds.), *Handbook of adolescent psychology* (pp. 15-54) Wiley Online Library.

Locuniak, M. N., & Jordan, N. C. (2008). Using kindergarten number sense to predict calculation fluency in second grade. *Journal of Learning Disabilities, 41*(5), 451-459. doi:10.1177/0022219408321126 [doi]

Lukowski, S. L., Soden, B., Hart, S. A., Thompson, L. A., Kovas, Y., & Petrill, S. A. (2014). Etiological distinction of working memory components in relation to mathematics. *Intelligence, 47*, 54-62. doi:10.1016/j.intell.2014.09.001 ER

Lummis, M., & Stevenson, H. W. (1990). Gender differences in beliefs and achievement - a cross-cultural-study. *Developmental Psychology, 26*(2), 254-263. doi:10.1037//0012-1649.26.2.254

Maheady, L., & Gard, J. (2010). Classwide peer tutoring: Practice, theory, research, and personal narrative. *Intervention in School and Clinic, 46*(2), 71-78. doi:10.1177/1053451210376359

Mammarella, N., Cornoldi, C., & Donadello, E. (2003). Visual but not spatial working memory deficit in children with spina bifida. *Brain and Cognition, 53*(2), 311-314.

Mariani, M. A., & Barkley, R. A. (1997). Neuropsychological and academic functioning in preschool boys with attention deficit hyperactivity disorder. *Developmental Neuropsychology, 13*(1), 111-129.

Martens, R., Hurks, P. P. M., Meijs, C., Wassenberg, R., & Jolles, J. (2011). Sex differences in arithmetical performance scores: Central tendency and variability. *Learning and Individual Differences, 21*(5), 549-554. doi:10.1016/j.lindif.2011.06.003

Martin, R. B., Cirino, P. T., Sharp, C., & Barnes, M. (2014). Number and counting skills in kindergarten as predictors of grade 1 mathematical skills. *Learning and Individual Differences, 34*, 12-23. doi:10.1016/j.lindif.2014.05.006

Martinussen, R., & Tannock, R. (2006). Working memory impairments in children with attention-deficit hyperactivity disorder with and without comorbid language learning disorders. *Journal of Clinical and Experimental Neuropsychology, 28*(7), 1073-1094. doi:10.1080/13803390500205700

165

Massetti, G. M., Lahey, B. B., Pelham, W. E., Loney, J., Ehrhardt, A., Lee, S. S., & Kipp, H. (2008). Academic achievement over 8 years among children who met modified criteria for attention-deficit/hyperactivity disorder at 4-6 years of age. *Journal of Abnormal Child Psychology, 36*(3), 399-410. doi:10.1007/s10802-007-9186-4

Mautone, J. A., DuPaul, G. J., & Jitendra, A. K. (2005). The effects of computer-assisted instruction on the mathematics performance and classroom behavior of children with ADHD. *Journal of Attention Disorders, 9*(1), 301-312. doi:9/1/301 [pii]

Mazzocco, M. M. M. (2007). Defining and differentiating mathematical learning disabilities and difficulties. In Berch, D.B. & Mazzocco, M.M.M (Ed.), *Why is math so hard for some children? the nature and origins of mathematical learning difficulties and disabilities* (pp. 29-47). Baltimore, MD: Brookes.

Mazzocco, M. M. M., Devlin, K. T., & McKenney, S. J. (2008). Is it a fact? timed arithmetic performance of children with mathematical learning disabilities (MLD) varies as a function of how MLD is defined. *Developmental Neuropsychology, 33*(3), 318-344. doi:10.1080/87565640801982403

McCallum, E., Skinner, C. H., & Hutchins, H. (2004). The taped-problems intervention: Increasing division fact fluency using a low-tech self-managed time-delay intervention. *Journal of Applied School Psychology, 20*(2), 129-147.

McClelland, M. M., Acock, A. C., & Morrison, F. J. (2006). The impact of kindergarten learning-related skills on academic trajectories at the end of elementary school. *Early Childhood Research Quarterly, 21*(4), 471-490.

McCloskey, M. (1992). Cognitive mechanisms in numerical processing - evidence from acquired dyscalculia. *Cognition, 44*(1-2), 107-157. doi:10.1016/0010-0277(92)90052-J

McCoach, D. B., & Yu, H. (2016). Using individual growth curves to model reading fluency. In K. D. Cummings, & Y. M. Petscher (Eds.), *The fluency construct: Curriculum-based measurement concepts and applications* (pp. 269-308). New York: Springer.

McDougall, D., & Brady, M. P. (1998). Initiating and fading self-management interventions to increase math fluency in general education classes. *Exceptional Children, 64*(2), 151-166.

McKenzie, B., Bull, R., & Gray, C. (2003). The effects of phonological and visual-spatial interference on children's arithmetical performance. *Educational and Child Psychology, 20*(3), 93-108.

McTiernan, A., Holloway, J., Healy, O., & Hogan, M. (2016). A randomized controlled trial of the morningside math facts curriculum on fluency, stability, endurance and application outcomes. *Journal of Behavioral Education, 25*(1), 49-68.

Melby-Lervag, M., & Hulme, C. (2013). Is working memory training effective? A meta-analytic review. *Developmental Psychology, 49*(2), 270-291. doi:10.1037/a0028228

Melby-Lervag, M., & Hulme, C. (2016). There is no convincing evidence that working memory training is effective: A reply to au et al. (2014) and karbach and verhaeghen (2014). *Psychonomic Bulletin & Review, 23*(1), 324-330. doi:10.3758/s13423-015-0862-z

Melby-Lervag, M., Redick, T. S., & Hulme, C. (2016). Working memory training does not improve performance on measures of intelligence or other measures of "far transfer": Evidence from a meta-analytic review. *Perspectives on Psychological Science : A Journal of the Association for Psychological Science, 11*(4), 512-34. doi:10.1177/1745691616635612

Melhuish, E. C., Phan, M. B., Sylva, K., Sammons, P., Siraj-Blatchford, I., & Taggart, B. (2008). Effects of the home learning environment and preschool center experience upon literacy and numeracy development in early primary school. *Journal of Social Issues, 64*(1), 95-114. doi:10.1111/j.1540-4560.2008.00550.x

Mercer, C., Mercer, K., & Campbell, K. (2002). Great leaps math. *Gainesville, FL: Diarmuid,*

Merrell, C., & Tymms, P. B. (2001). Inattention, hyperactivity and impulsiveness: Their impact on
    academic achievement and progress. *British Journal of Educational Psychology, 71*, 43-56.
    doi:10.1348/000709901158389

Metcalfe, A. W. S., Ashkenazi, S., Rosenberg-Lee, M., & Menon, V. (2013). Fractionating the neural
    correlates of individual working memory components underlying arithmetic problem solving skills
    in children. *Developmental Cognitive Neuroscience, 6*, 162-175. doi:10.1016/j.dcn.2013.10.001

Metcalfe, L. A., Harvey, E. A., & Laws, H. B. (2013). The longitudinal relation between
    academic/cognitive skills and externalizing behavior problems in preschool children. *Journal of
    Educational Psychology, 105*(3), 881-894. doi:10.1037/a0032624

Methe, S. A., Kilgus, S. P., Neiman, C., & Riley-Tillman, T. C. (2012). Meta-analysis of interventions for
    basic mathematics computation in single-case research. *Journal of Behavioral Education, 21*(3),
    230-253.

Methe, S., Briesch, A., & Hulac, D. (2015). Evaluating procedures for reducing measurement error in
    math curriculum-based measurement probes. *Assessment for Effective Intervention, 40*(2), 99-113.
    doi:10.1177/1534508414553295

Meyers, J., & Nastasi, B. K. (1999). Primary prevention in school settings. In C. R. Reynolds, & T. B.
    Gutkin (Eds.), *The handbook of school psychology.* (3rd ed., pp. 764-799). New York, NY: Wiley.

Miyake, A., & Shah, P. (Eds.). (1999). *Models of working memory: Mechanisms of active maintenance
    and executive control* . Cambridge, UK: Cambridge University Press.

Miyake, A., Friedman, N., Rettinger, D., Shah, P., & Hegarty, P. (2001). How are visuospatial working
    memory, executive functioning, and spatial abilities related? A latent-variable analysis. *Journal of
    Experimental Psychology-General, 130*(4), 621-640. doi:10.1037/0096-3445.130.4.621

Mok, M. M. C., McInerney, D. M., Zhu, J., & Or, A. (2015). Growth trajectories of mathematics
achievement: Longitudinal tracking of student academic progress. *British Journal of Educational Psychology, 85*(2), 154-171. doi:10.1111/bjep.12060

Mong, M. D., & Mong, K. W. (2010). Efficacy of two mathematics interventions for enhancing fluency
with elementary students. *Journal of Behavioral Education, 19*(4), 273-288.

Montori, V. M., & Rothman, R. L. (2005). Weakness in numbers - the challenge of numeracy in health
care. *Journal of General Internal Medicine, 20*(11), 1071-1072. doi:10.1111/j.1525-1497.2005.051498.x

Morgan, P. L., Farkas, G., & Wu, Q. (2009). Five-year growth trajectories of kindergarten children with
learning difficulties in mathematics. *Journal of Learning Disabilities, 42*(4), 306-321.
doi:10.1177/0022219408331037

Morgan, P. L., Farkas, G., & Wu, Q. (2011). Kindergarten children's growth trajectories in reading and
mathematics: Who falls increasingly behind? *Journal of Learning Disabilities, 44*(5), 472-488.
doi:10.1177/0022219411414010

Murayama, K., Pekrun, R., Lichtenfeld, S., & Vom Hofe, R. (2013). Predicting long-term growth in
students' mathematics achievement: The unique contributions of motivation and cognitive strategies.
*Child Development, 84*(4), 1475-1490.

Muthen, B. O., & Khoo, S. T. (1998). Longitudinal studies of achievement growth using latent variable
modeling. *Learning and Individual Differences, 10*(2), 73-101. doi:10.1016/S1041-6080(99)80135-6

National Council of Teachers of Mathematics. (2000). In Commission on Standards for School
Mathematics (Ed.), *Curriculum and evaluation standards for school mathematics.* Reston, VA:
Author.

National Mathematics Advisory Panel. (2008). In Author (Ed.), *The final report of the national mathematics advisory panel.* U.S. Department of Education.

National Summer Learning Association. (2013). Teachers confirm time wasted due to summer learning loss, 25 national summer learning association new vision for summer school districts working to close achievement gap . Retrieved from https://c.ymcdn.com/sites/www.summerlearning.org/resource/resmgr/press_releases/nsla_summer_release_130528_f.pdf

Nelson, P. M., Parker, D. C., & Zaslofsky, A. F. (2016). The relative value of growth in math fact skills across late elementary and middle school. *Assessment for Effective Intervention, 41*(3), 184-192.

Nelson, P. M., Burns, M. K., Kanive, R., & Ysseldyke, J. E. (2013). Comparison of a math fact rehearsal and a mnemonic strategy approach for improving math fact fluency. *Journal of School Psychology, 51*(6), 659-667. doi:10.1016/j.jsp.2013.08.003

Nesselroade, J. R., & Baltes, P. B. (1979). *Longitudinal research in the study of behavior and development.* New York, NY: Academic Press.

Nezlek, J. B. (2008). An introduction to multilevel modeling for social and personality psychology. *Social and Personality Psychology Compass, 2*(2), 842-860.

Ontario Ministry of Education. (2005). *The ontario curriculum, grades 1-8*. Retrieved from http://www.edu.gov.on.ca/eng/curriculum/elementary/math18curr.pdf

Ostad, S. A. (1997). Developmental differences in addition strategies: A comparison of mathematically disabled and mathematically normal children. *British Journal of Educational Psychology, 67*, 345-357.

Paechter, M., Luttenberger, S., Macher, D., Berding, F., Papousek, I., Weiss, E. M., & Fink, A. (2015). The effects of nine-week summer vacation: Losses in mathematics and gains in reading. *Eurasia Journal of Mathematics Science and Technology Education, 11*(6), 1399-1413.

Pagani, L. S., Fitzpatrick, C., Archambault, I., & Janosz, M. (2010). School readiness and later achievement: A french canadian replication and extension. *Developmental Psychology, 46*(5), 984-994. doi:10.1037/a0018881

Plamondon, A., & Martinussen, R. (2015). Inattention symptoms are associated with academic achievement mostly through variance shared with intrinsic motivation and behavioral engagement. Journal of Attention Disorders, 1-13. DOI: 10.1177/1087054715587098

Pariseau, M. E., Fabiano, G. A., Massetti, G. M., Hart, K. C., & Pelham, W. E., Jr. (2010). Extended time on academic assignments: Does increased time lead to improved performance for children with attention-deficit/hyperactivity disorder? *School Psychology Quarterly, 25*(4), 236-248. doi:10.1037/a0022045

Parsons, S., & Bryner, J. (2005). *Does numercay matter more?*. London: National Research and Development Centre for Adult Literary and Numeracy.

Passolunghi, M. C., & Cornoldi, C. (2008). Working memory failures in children with arithmetical difficulties. *Child Neuropsychology, 14*(5), 387-400.

Passolunghi, M. C., & Siegel, L. S. (2004). Working memory and access to numerical information in children with disability in mathematics. *Journal of Experimental Child Psychology, 88*(4), 348-367. doi:10.1016/j.jecp.2004.04.002

Passolunghi, M. C., & Costa, H. M. (2016). Working memory and early numeracy training in preschool children. *Child Neuropsychology, 22*(1), 81-98. doi:10.1080/09297049.2014.971726

Passolunghi, M. C., & Mammarella, I. C. (2010). Spatial and visual working memory ability in children

    with difficulties in arithmetic word problem solving. *European Journal of Cognitive Psychology,*

    *22*(6), 944-963. doi:10.1080/09541440903091127

Patton, K. L. S., & Reschly, A. L. (2013). Using curriculum-based measurement to examine summer

    learning loss. *Psychology in the Schools, 50*(7), 738-753. doi:10.1002/pits.21704

Pelham, W., Gnagy, E., Greiner, A., Hoza, B., Hinshaw, S., Swanson, J., . . . McBurnett, K. (2000).

    Behavioral versus behavioral and pharmacological treatment in ADHD children attending a summer

    treatment program. *Journal of Abnormal Child Psychology, 28*(6), 507-525.

    doi:10.1023/A:1005127030251

Peng, P., & Fuchs, D. (2014). A Meta-

    Analysis of Working Memory Deficits in Children With Learning Difficulties: Is there

    a Difference BetweenVerbal Domain and Numerical Domain? *Journal of Learning Disabilities, ,* 1-

    18.

Peng, P., Namkung, J., Barnes, M., & Sun, C. (2016). A meta-analysis of mathematics and working

    memory: Moderating effects of working memory domain, type of mathematics skill, and sample

    characteristics. *Journal of Educational Psychology, 108*(4), 455-473. doi:10.1037/edu0000079

Petrill, S., Logan, J., Hart, S., Vincent, P., Thompson, L., Kovas, Y., & Plomin, R. (2012). Math fluency

    is etiologically distinct from untimed math performance, decoding fluency, and untimed reading

    performance: Evidence from a twin study. *Journal of Learning Disabilities, 45*(4), 371-381.

    doi:10.1177/0022219411407926

Phillips, L. M., Norris, S. P., Osmond, W. C., & Maynard, A. M. (2002). Relative reading achievement: A

    longitudinal study of 187 children from first through sixth grades. *Journal of Educational*

    *Psychology, 94*(1), 3-13. doi:10.1037//0022-0663.94.1.3

Pingault, J., Cote, S. M., Vitaro, F., Falissard, B., Genolini, C., & Tremblay, R. E. (2014). The

developmental course of childhood inattention symptoms uniquely predicts educational attainment:

A 16-year longitudinal study. *Psychiatry Research, 219*(3), 707-709.

doi:10.1016/j.psychres.2014.06.022

Pingault, J., Tremblay, R. E., Vitaro, F., Carbonneau, R., Genolini, C., Falissard, B., & Cote, S. M.

(2011). Childhood trajectories of inattention and hyperactivity and prediction of educational

attainment in early adulthood: A 16-year longitudinal population-based study. *American Journal of

Psychiatry, 168*(11), 1164-1170. doi:10.1176/appi.ajp.2011.10121732

Plass, J. L., O'Keefe, P. A., Homer, B. D., Case, J., Hayward, E. O., Stein, M., & Perlin, K. (2013). The

impact of individual, competitive, and collaborative mathematics game play on learning,

performance, and motivation. *Journal of Educational Psychology, 105*(4), 1050-1066.

Polderman, T. J., Derks, E. M., Hudziak, J. J., Verhulst, F. C., Posthuma, D., & Boomsma, D. I. (2007).

Across the continuum of attention skills: A twin study of the SWAN ADHD rating scale. *Journal of

Child Psychology and Psychiatry, 48*(11), 1080-1087.

Polderman, T. J. C., Boomsma, D. I., Bartels, M., Verhulst, F. C., & Huizink, A. C. (2010). A systematic

review of prospective studies on attention problems and academic achievement. *Acta Psychiatrica

Scandinavica, 122*(4), 271-284. doi:10.1111/j.1600-0447.2010.01568.x

Poncy, B. C., Skinner, C. H., & O'Mara, T. (2006). Detect, practice, and repair: The effects of a classwide

intervention on elementary students' math-fact fluency. *Journal of Evidence-Based Practices for

Schools,*

Poncy, B. C., Skinner, C. H., & McCallum, E. (2012). A comparison of class-wide taped problems and

cover, copy, and compare for enhancing mathematics fluency. *Psychology in the Schools, 49*(8),

744-755. doi:10.1002/pits.21631

Poncy, B. C., Solomon, B., Duhon, G., Skinner, C., Moore, K., & Simons, S. (2015). An analysis of

    learning rate and curricular scope: Caution when choosing academic interventions based on

    aggregated outcomes. *School Psychology Review, 44*(3), 289-305.

Poncy, B., Duhon, G., Lee, S., & Key, A. (2010). Evaluation of techniques to promote generalization with

    basic math fact skills. *Journal of Behavioral Education, 19*(1), 76-92. doi:10.1007/s10864-010-

    9101-x

Prado, J., Mutreja, R., & Booth, J. R. (2014). Developmental dissociation in the neural responses to

    simple multiplication and subtraction problems. *Developmental Science, 17*(4), 537-552.

    doi:10.1111/desc.12140

Price, G. R., Mazzocco, M. M. M., & Ansari, D. (2013). Why mental arithmetic counts: Brain activation

    during single digit arithmetic predicts high school math scores. *Journal of Neuroscience, 33*(1), 156-

    163. doi:10.1523/JNEUROSCI.2936-12.2013

Prindle, J., Mitchell, A. M., & Pretscher, Y. (2016). Using response time and accuracy data to inform the

    measurement of fluency. In K. D. Cummings, & Y. M. Petscher (Eds.), *The fluency construct:*

    *Curriculum-based measurement concepts and applications* (pp. 165-186). New York: Springer.

Protopapas, A., Sideridis, G. D., Mouzaki, A., & Simos, P. G. (2011). Matthew effects in reading

    comprehension: Myth or reality? *Journal of Learning Disabilities, 44*(5), 402-420.

    doi:10.1177/0022219411417568

Rabiner, D. L., Murray, D. W., Schmid, L., & Malone, P. S. (2004). An exploration of the relationship

    between ethnicity, attention problems, and academic achievement. *School Psychology Review, 33*(4),

    498.

Rabiner, D. L., Murray, D. W., Rosen, L., Hardy, K., Skinner, A., & Underwood, M. (2010). Instability in

    teacher ratings of children's inattentive symptoms: Implications for the assessment of ADHD.

*Journal of Developmental and Behavioral Pediatrics : JDBP, 31*(3), 175-180.

doi:10.1097/DBP.0b013e3181d5a2d8 [doi]

Rabiner, D. L., Carrig, M. M., & Dodge, K. A. (2016). Attention problems and academic achievement:

Do persistent and earlier-emerging problems have more adverse long-term effects? *Journal of*

*Attention Disorders, 20*(11), 946-957. doi:10.1177/1087054713507974

Raggi, V. L., & Chronis, A. M. (2006). Interventions to address the academic impairment of children and

adolescents with ADHD. *Clinical Child and Family Psychology Review, 9*(2), 85-111.

doi:10.1007/s10567-006-0006-0

Raghubar, K. P., Barnes, M. A., & Hecht, S. A. (2010). Working memory and mathematics: A review of

developmental, individual difference, and cognitive approaches. *Learning and Individual*

*Differences, 20*(2), 110-122. doi:10.1016/j.lindif.2009.10.005

Raghubar, K., Cirino, P., Barnes, M., Ewing-Cobbs, L., Fletcher, J., & Fuchs, L. (2009). Errors in multi-

digit arithmetic and behavioral inattention in children with math difficulties. *Journal of Learning*

*Disabilities, 42*(4), 356-371. doi:10.1177/0022219409335211

Ramani, G. B., & Siegler, R. S. (2008). Promoting broad and stable improvements in low-income

children's numerical knowledge through playing number board games. *Child Development, 79*(2),

375-394.

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*(6), 543-549. doi:10.1007/s10643-008-0234-7

Rasmussen, C., & Bisanz, J. (2005). Representation and working memory in early arithmetic. *Journal of*

*Experimental Child Psychology, 91*(2), 137-157. doi:10.1016/j.jecp.2005.01.004

Raudenbush, S. W., Bryk, A. S., & Congdon, R. (2010). *HLM 7.00 for windows*. Lincolnwood, IL: Scientific Software International Inc.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* Sage.

Raudenbush, S., Bryk, A., Cheong, Y., Congdon, R., & Du Toit, M. (2004). *HLM 6: Hierarchical linear and nonlinear modeling.* Lincolnwood, IL: Scientific Software International.

Ready, D. D. (2013). Associations between student achievement and student learning: Implications for value-added school accountability models. *Educational Policy, 27*(1), 92-120. doi:10.1177/0895904811429289

Redick, T., Shipstead, Z., Wiemers, E., Melby-Lervåg, M., & Hulme, C. (2015). What's working in working memory training? an educational perspective. *Educational Psychology Review, 27*(4), 617-633. doi:10.1007/s10648-015-9314-6

Reid, R., Trout, A. L., & Schartz, M. (2005). Self-regulation interventions for children with attention deficit/hyperactivity disorder. *Exceptional Children, 71*(4), 361.

Reisener, C. D., Dufrene, B. A., Clark, C. R., Olmi, D. J., & Tingstrom, D. H. (2016). Selecting effective interventions to increase math computation fluency via brief experimental analyses. *Psychology in the Schools, 53*(1), 39-57. doi:10.1002/pits.21887

Rennie, B., Beebe-Frankenberger, M., & Swanson, H. L. (2014). A longitudinal study of neuropsychological functioning and academic achievement in children with and without signs of attention-deficit/hyperactivity disorder. *Journal of Clinical and Experimental Neuropsychology, 36*(6), 621-635. doi:10.1080/13803395.2014.921284

Rescorla, L., & Rosenthal, A. S. (2004). Growth in standardized ability and achievement test scores from 3rd to 10th grade. *Journal of Educational Psychology, 96*(1), 85-96. doi:10.1037/0022-0663.96.1.85

Reyna, V. F., & Brainerd, C. J. (2007). The importance of mathematics in health and human judgment: Numeracy, risk communication, and medical decision making. *Learning and Individual Differences, 17*(2), 147-159. doi:10.1016/j.lindif.2007.03.010

Reys, B., Reys, R., & Koyama, M. (1996). The development of computation in three japanese primary-grade textbooks. *Elementary School Journal, 96*(4), 423-437. doi:10.1086/461837

Rhymer, K. N., Dittmer, K. I., Skinner, C. H., & Jackson, B. (2000). Effectiveness of a multi-component treatment for improving mathematics fluency. *School Psychology Quarterly, 15*(1), 40.

Rhymer, K. N., Skinner, C. H., Jackson, S., McNeill, S., Smith, T., & Jackson, B. (2002). The I-minute explicit timing intervention: The influence of mathematics problem difficulty. *Journal of Instructional Psychology, 29*(4), 305.

Rittle-Johnson, B., Siegler, R., & Alibali, M. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process. *Journal of Educational Psychology, 93*(2), 346-362. doi:10.1037/0022-0663.93.2.346

Rittle-Johnson, B., Schneider, M., & Star, J. R. (2015). Not a one-way street: Bidirectional relations between procedural and conceptual knowledge of mathematics. *Educational Psychology Review, 27*(4), 587-597. doi:10.1007/s10648-015-9302-x

Rivera, D., & Bryant, B. (1992). Mathematics instruction for students with special needs. *Intervention in School and Clinic, 28*(2), 71-86. doi:10.1177/105345129202800203

Rivera-Batiz, F. L. (1992). Quantitative literacy and the likelihood of employment among young-adults in the united-states. *Journal of Human Resources, 27*(2), 313-328. doi:10.2307/145737

Roberts, G., Quach, J., Spencer-Smith, M., Anderson, P. J., Gathercole, S., Gold, L., . . . Wake, M. (2016). Academic outcomes 2 years after working memory training for children with low working

memory A randomized clinical trial. *Jama Pediatrics, 170*(5), e154568.
doi:10.1001/jamapediatrics.2015.4568

Robinson, K. M., & Dubé, A. K. (2013). Children's additive concepts: Promoting understanding and the
role of inhibition. *Learning and Individual Differences, 23*, 101-107.

Robinson, C. S., Menchetti, B. M., & Torgesen, J. K. (2002). Toward a two-factor theory of one type of
mathematics disabilities. *Learning Disabilities Research & Practice, 17*(2), 81-89.
doi:10.1111/1540-5826.00035

Robinson, K. M. (2001). The validity of verbal reports in children's subtraction. *Journal of Educational
Psychology, 93*(1), 211-222. doi:10.1037/0022-0663.93.1.211

Rodriguez, A., Jaervelin, M., Obel, C., Taanila, A., Miettunen, J., Moilanen, I., . . . Olsen, J. (2007). Do
inattention and hyperactivity symptoms equal scholastic impairment? evidence from three european
cohorts. *Bmc Public Health, 7*, 327. doi:10.1186/1471-2458-7-327

Rogers, M., Hwang, H., Toplak, M., Weiss, M., & Tannock, R. (2011). Inattention, working memory, and
academic achievement in adolescents referred for attention deficit/hyperactivity disorder (ADHD).
*Child Neuropsychology, 17*(5), 444-458. doi:10.1080/09297049.2010.544648

Ross, J., Hogaboam-Gray, A., McDougall, D., & Bruce, C. (2002). The contribution of technology to
mathematics education reform.

Rotzer, S., Loenneker, T., Kucian, K., Martin, E., Klaver, P., & von Aster, M. (2009). Dysfunctional
neural network of spatial working memory contributes to developmental dyscalculia.
*Neuropsychologia, 47*(13), 2859-2865. doi:10.1016/j.neuropsychologia.2009.06.009

Royer, J. M., Tronsky, L. N., Chan, Y., Jackson, S. J., & Marchant, H. (1999). Math-fact retrieval as the
cognitive mechanism underlying gender differences in math test performance. *Contemporary
Educational Psychology, 24*(3), 181-266. doi:10.1006/ceps.1999.1004

178

Salaschek, M., Zeuch, N., & Souvignier, E. (2014). Mathematics growth trajectories in first grade:

Cumulative vs. compensatory patterns and the role of number sense. *Learning and Individual*

*Differences, 35*, 103-112. doi:10.1016/j.lindif.2014.06.009

Scarborough, H. S., & Parker, J. D. (2003). Matthew effects in children with learning disabilities:

Development of reading, IQ, and psychosocial problems from grade 2 to grade 8. *Annals of*

*Dyslexia, 53*(1), 47-71.

Schulte, A. C., & Stevens, J. J. (2015). Once, sometimes, or always in special education: Mathematics

growth and achievement gaps. *Exceptional Children, 81*(3), 370-387.

doi:10.1177/0014402914563695

Schutte, G. M., Duhon, G. J., Solomon, B. G., Poncy, B. C., Moore, K., & Story, B. (2015). A

comparative analysis of massed vs. distributed practice on basic math fact fluency growth rates.

*Journal of School Psychology, 53*(2), 149-159. doi:10.1016/j.jsp.2014.12.003

Shanley, L. (2016). Evaluating longitudinal mathematics achievement growth: Modeling and

measurement considerations for assessing academic progress. *Educational Researcher,*

doi:10.3102/0013189X16662461

Shapiro, E. S., DuPaul, G. J., Bradley, K. L., & Bailey, L. T. (1996). A school-based consultation

program for service delivery to middle school students with attention-deficit/hyperactivity disorder.

*Journal of Emotional and Behavioral Disorders, 4*(2), 73-81.

Shin, J., Espin, C. A., Deno, S. L., & McConnell, S. (2004). Use of hierarchical linear modeling and

curriculum-based measurement for assessing academic growth and instructional factors for students

with learning difficulties. *Asia Pacific Education Review, 5*(2), 136-148.

Shin, T., Davison, M. L., Long, J. D., Chan, C., & Heistad, D. (2013). Exploring gains in reading and

    mathematics achievement among regular and exceptional students using growth curve modeling.

    *Learning and Individual Differences, 23*, 92-100. doi:10.1016/j.lindif.2012.10.002

Shrager, J., & Siegler, R. S. (1998). SCADS: A model of children's strategy choices and strategy

    discoveries. *Psychological Science, 9*(5), 405-410. doi:10.1111/1467-9280.00076

Siegler, R. S. (1987). Strategy choices in subtraction. In J. A. Sloboda, & D. Rogers (Eds.), *Cognitive*

    *processes in mathematics* (). New York, NY: Oxford University Press.

Siegler, R. S., & Shipley, C. (1995). Variation, selection, and cognitive change. In T. Simon, & G.

    Halford (Eds.), *Developing cognitive competence: New approaches to process modeling.* (pp. 31-

    76). Hillsdale, NJ: Erlbaum.

Siegler, R. S., & Shrager, J. (1984). Strategy choices in addition and subtraction: How do children know

    what to do? In C. Sophia (Ed.), *The origins of cognitive skills.* (pp. 229-293). Hillsdale, NJ: Erlbaum.

Siegler, R. S., & Stern, E. (1998). Conscious and unconscious strategy discoveries: A microgenetic

    analysis. *Journal of Experimental Psychology-General, 127*(4), 377-397. doi:10.1037/0096-

    3445.127.4.377

Siegler, R. S. (2009). Improving the numerical understanding of children from low-income families.

    *Child Development Perspectives, 3*(2), 118-124.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event*

    *occurrence* Oxford university press.

Singer-Dudek, J., & Greer, R. D. (2005). A long-term analysis of the relationship between fluency and the

    training and maintenance of complex math skills. *Psychological Record, 55*(3), 361-376.

Skinner, C. H., McLaughlin, T., & Logan, P. (1997). Cover, copy, and compare: A self-managed academic intervention effective across skills, students, and settings. *Journal of Behavioral Education, 7*(3), 295-306.

Skinner, C. H., Turco, T. L., Beatty, K. L., & Rasavage, C. (1989). Cover, copy, and compare: A method for increasing multiplication performance. *School Psychology Review,*

Sowinski, C., LeFevre, J., Skwarchuk, S., Kamawar, D., Bisanz, J., & Smith-Chant, B. (2015). Refining the quantitative pathway of the pathways to mathematics model. *Journal of Experimental Child Psychology, 131*, 73-93. doi:10.1016/j.jecp.2014.11.004

Speelman, C. P., & Kirsner, K. (2005). Skill acquisition: History, questions, and theories. *Beyond the learning curve: The construction of mind* (pp. 26-64) Oxford University Press on Demand.

Stanovich, K. E. (1986). Matthew effects in reading - some consequences of individual-differences in the acquisition of literacy. *Reading Research Quarterly, 21*(4), 360-407. doi:10.1598/RRQ.21.4.1

Stevens, J. J., Schulte, A. C., Elliott, S. N., Nese, J. F. T., & Tindal, G. (2015). Growth and gaps in mathematics achievement of students with and without disabilities on a statewide achievement test. *Journal of School Psychology, 53*(1), 45-62. doi:10.1016/j.jsp.2014.11.001

Stipek, D., & Valentino, R. A. (2015). Early childhood memory and attention as predictors of academic growth trajectories. *Journal of Educational Psychology, 107*(3), 771-788. doi:10.1037/edu0000004

Stocker, J. D., & Kubina, R. M. (2016). Impact of cover, copy, and compare on fluency outcomes for students with disabilities and math deficits: A review of the literature. *Preventing School Failure: Alternative Education for Children and Youth, ,* 1-10. doi:10.1080/1045988X.2016.1196643

Swanson, J., Schuck, S., Mann, M., Carlson, C., Hartman, K. & Joe, S. (2004). Categorical and dimensional definitions and evaluations of symptoms of ADHD: The SNAP and the SWAN ratings scales. Retrieved from www.adhd.net

Swanson, J. M., Schuck, S., Porter, M. M., Carlson, C., Hartman, C. A., Sergeant, J. A., . . . Lakes, K. (2012). Categorical and dimensional definitions and evaluations of symptoms of ADHD: History of the SNAP and the SWAN rating scales. *The International Journal of Educational and Psychological Assessment, 10*(1), 51.

Swanson, H. L., & Sachse-Lee, C. (2000). A meta-analysis of single-subject-design intervention research for students with LD. *Journal of Learning Disabilities, 33*(2), 114-136.

Swanson, H. L. (2011). Working memory, attention, and mathematical problem solving: A longitudinal study of elementary school children. *Journal of Educational Psychology, 103*(4), 821-837. doi:10.1037/a0025114

Tenison, C., & Anderson, J. R. (2015). Modeling the distinct phases of skill acquisition. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* doi:http://dx.doi.org.myaccess.library.utoronto.ca/10.1037/xlm0000204

Thevenot, C., Barrouillet, P., Castel, C., & Uittenhove, K. (2016). Ten-year-old children strategies in mental addition: A counting model account. *Cognition, 146*, 48-57. doi:10.1016/j.cognition.2015.09.003

Thorell, L. B. (2007). Do delay aversion and executive function deficits make distinct contributions to the functional impact of ADHD symptoms? A study of early academic skill deficits. *Journal of Child Psychology and Psychiatry, 48*(11), 1061-1070.

Thurber, R. S., Shinn, M. R., & Smolkowski, K. (2002). What is measured in mathematics tests? construct validity of curriculum-based mathematics measures. *School Psychology Review, 31*(4), 498-513.

Titz, C., & Karbach, J. (2014). Working memory and executive functions: Effects of training on academic achievement. *Psychological Research, 78*(6), 852-868. doi:10.1007/s00426-013-0537-1

Tucker, J. A. (1985). Curriculum-based assessment: An introduction. *Exceptional Children,*

Vale, C., Weaven, M., Davies, A., Hooley, N., Davidson, K., & Loton, D. (2013). Growth in literacy and numeracy achievement: Evidence and explanations of a summer slowdown in low socio-economic schools. *Australian Educational Researcher, 40*(1), 1-25. doi:10.1007/s13384-012-0065-9

Van de Weijer-Bergsma, E., Kroesbergen, E. H., & Van Luit, J. E. H. (2015). Verbal and visual-spatial working memory and mathematical ability in different domains throughout primary school. *Memory & Cognition, 43*(3), 367-378.

Vanbinst, K., Ceulemans, E., Ghesquiere, P., & De Smedt, B. (2015). Profiles of children's arithmetic fact development: A model-based clustering approach. *Journal of Experimental Child Psychology, 133*, 29-46. doi:10.1016/j.jecp.2015.01.003

Vanbinst, K., Ghesquiere, P., & De Smedt, B. (2014). Arithmetic strategy development and its domain-specific and domain-general cognitive correlates: A longitudinal study in children with persistent mathematical learning difficulties. *Research in Developmental Disabilities, 35*(11), 3001-3013. doi:10.1016/j.ridd.2014.06.023 ER

Vasilyeva, M., Laski, E. V., & Shen, C. (2015). Computational fluency and strategy choice predict individual and cross-national differences in complex arithmetic. *Developmental Psychology, 51*(10), 1489-1500. doi:10.1037/dev0000045

Verachtert, P., Van Damme, J., Onghena, P., & Ghesquiere, P. (2009). A seasonal perspective on school effectiveness: Evidence from a flemish longitudinal study in kindergarten and first grade. *School Effectiveness and School Improvement, 20*(2), 215-233. doi:10.1080/09243450902883896

Villasenor, A., & Kepner, H. S. (1993). Arithmetic from a problem-solving perspective - an urban implementation. *Journal for Research in Mathematics Education, 24*(1), 62-69. doi:10.2307/749386

Walberg, H. J., & Tsai, S. (1983). Matthew effects in education. *American Educational Research Journal, 20*(3), 359-373.

Wechsler, D. (2003). Wechsler intelligence scale for children–Fourth edition (WISC-IV). *San Antonio, TX: The Psychological Corporation,*

Whitney, T., Hirn, R. G., & Lingo, A. S. (2016). Effects of a mathematics fluency program on mathematics performance of students with challenging behaviors. *Preventing School Failure: Alternative Education for Children and Youth, 60*(2), 133-142. doi:10.1080/1045988X.2015.1040984

Willcutt, E. G. (2012). The prevalence of DSM-IV attention-deficit/hyperactivity disorder: A meta-analytic review. *Neurotherapeutics, 9*(3), 490-499. doi:10.1007/s13311-012-0135-8

Wintre, M. G. (1986). Challenging the assumption of generalized academic losses over summer. *Journal of Educational Research, 79*(5), 308-312.

Woltman, H., Feldstain, A., MacKay, J. C., & Rocchi, M. (2012). An introduction to hierarchical linear modeling. *Tutorials in Quantitative Methods for Psychology, 8*(1), 52-69.

Ysseldyke, J. E., Vanderwood, M. L., & Shriner, J. (1997). Changes over the past decade in special education referral to placement probability an incredibly reliable practice. *Assessment for Effective Intervention, 23*(1), 193-201.

Ysseldyke, J., Thill, T., Pohl, J., & Bolt, D. (2005). Using MathFacts in a flash to enhance computational fluency. *Journal of Evidence-Based Practices for Schools, 6*(1), 59-89.

Zentall, S. S. (1990). Fact-retrieval automatization and math problem-solving by learning-disabled, attention-disordered, and normal adolescents. *Journal of Educational Psychology, 82*(4), 856-865. doi:10.1037/0022-0663.82.4.856

Zentall, S. S., & Smith, Y. N. (1993). Mathematical performance and behavior of children with

    hyperactivity with and without coexisting aggression. *Behaviour Research and Therapy, 31*(7), 701-

    710. doi:10.1016/0005-7967(93)90123-C

Zentall, S. S., Smith, Y. N., Lee, Y. B. B., & Wieczorek, C. (1994). Mathematical outcomes of attention-

    deficit hyperactivity disorder. *Journal of Learning Disabilities, 27*(8), 510-519.

Zorzi, M., Priftis, K., & Umilta, C. (2002). Brain damage - neglect disrupts the mental number line.

    *Nature, 417*(6885), 138-139. doi:10.1038/417138a