Title
The Effects of Executive Functioning, Demographics, and School Factors on Mathematics Achievement Growth During Elementary School: A Multilevel, Multivariate, and Longitudinal Analysis

Author
Gundogdu, Mahmut

Publication Date
2019

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA
RIVERSIDE

The Effects of Executive Functioning, Demographics, and School Factors on Mathematics Achievement Growth During Elementary School: A Multilevel, Multivariate, and Longitudinal Analysis

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

Doctor of Philosophy

in

Education

by

Mahmut Gundogdu

June 2019

Dissertation Committee:
Dr. Lee Swanson, Co-Chairperson
Dr. Gregory J. Palardy, Co-Chairperson
Dr. Rollanda O'Connor

The Dissertation of Mahmut Gundogdu is approved:

Committee Co-Chairperson

Committee Co-Chairperson

University of California, Riverside

Dedication

I would like to dedicate this dissertation to my lovely wife and my son.

ABSTRACT OF THE DISSERTATION

The Effects of Executive Functioning, Demographics, and School Factors on Mathematics Achievement Growth During Elementary School: A Multilevel, Multivariate, and Longitudinal Analysis

by

Mahmut Gundogdu

Doctor of Philosophy, Graduate Program in Education
University of California, Riverside, June 2019
Dr. Lee Swanson, Co-Chairperson
Dr. Gregory J. Palardy, Co-Chairperson

This study examines how gains in mathematics achievement are related to executive processing functions and student sociodemographic characteristics across schools' national representative longitudinal sample of children in kindergarten (K) followed through grade four in the Early Childhood Longitudinal Study of 2010.

Mathematics trajectories were nonlinear, with greater gains in early versus later grades and small drops each summer. Children entering K with lower math demonstrated steeper gains over time. Relative to Caucasian children, Hispanic and African American children entered K with lower math. Hispanic children had higher growth rates whereas African American children had lower growth rates. Girls entered K with higher math, but boys gained more over time. Lower SES was associated with lower math but also steeper increases. Demographic factors explained a larger proportion of between-school differences in mathematics achievement than within-school differences in initial levels and long-term gains but not summer drops (66.23% versus 8.11% of variance at K; 17.02% versus 4.55% of variance in gains; 20% versus 0% of variance in summer

drops). Similarly, shaped trajectories and demographic effects were found for working memory, while the cognitive flexibility trajectory (measured only from 2nd grade) was a more linear. Critically, just as lower math scores in K were associated with steeper growth, lower working memory in K was also associated with steeper trajectories. In addition, positive associations were observed between working memory and math in K, trajectories of working memory and cognitive flexibility were strongly associated with trajectories of mathematics achievement. A similar pattern was observed for cognitive flexibility. These inclusions of demographic covariates did not alter these associations. Overall, these findings bolster the independent importance of executive function and sociodemographic factors, with the latter explaining a large amount of between-school variability. Education stakeholders such as teachers, school administrations, and school district can rely on the research findings in designing practical models of teaching mathematics that will take into consideration the role of sociodemographic factors as well as executive functioning among students, thereby improving their overall performance.

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

The purpose of this chapter is to discuss the importance of understanding the role of Executive Function (EF) on mathematics achievement when critical demographic characteristics are accounted for over a five-year period during early childhood education. This will be discussed in the context of the present literature. The theoretical perspectives to best understand this investigation are the comparisons between learning models that describe the trajectory of academic achievement when considering variances that result from demographic characteristics. The concern for mathematics achievement is not limited to educators, education researchers and policymakers receiving media attention. Concluding the chapter will be an explanation of the need for this investigation in the context of the present literature, how this investigation will contribute to the current conversation of mathematics achievement, the importance of the project, and the proposed research questions that will make contribute to the literature.

The Influence of Executive Function (EF) on Mathematics Achievement

Mathematics achievement is influenced by a host of cognitive processing functions that are in turn influenced by the learning opportunities made available through demographic characteristics. This investigation aims to study the influence of these variables using a longitudinal design to account for the nested nature of the educational system. It uses data collected during the critical period of early childhood education. Executive function is reported as the center for cognitive processing and is composed of a series of sophisticated and well-coordinated processing functions. Thus, a bottom-up review of the literature will be used to introduce the subject here briefly. EF is credited with facilitating the function of a myriad of basic tasks related to mathematics achievement. This central processing system is also critical in understanding more complex processes such as academic growth during a child's early education when the executive function skills are developing and the child is learning basic reading and mathematics skills (Diamond, 2013; Geary, 2013; Swanson & Alloway, 2012). Understanding the processes associated with EF and how they influence problem solving and academic achievement has been the focus of education researchers. Therefore, it is important to use a different methodology to replicate and better understand the previous findings between the relationship between EF and mathematics achievement. To do so requires understanding sources of variability between individuals, such as demographic characteristics, that are critical to understanding the findings reported in the literature. This investigation plans to expand on previous findings by closely considering the relationships between the different factors that influence mathematics achievement over time, during the critical period of the early childhood education.

In a review of the literature, Zelazo (2015) pointed out that, at a basic level, EF facilitates the processes related to academic achievements such as attention and behavior. Using the iterative reprocessing (IR) model as a reference, Zelazo (2015) pointed out that EF refers to neurocognitive skills that researchers have deemed important for aspects of everyday functioning. That is, any task that is successfully completed, from counting and reading a book, to long-term goals (e.g. career goals), has been attributed to EF successfully filling its role. More specifically, these central processes are critical in directing an individual's basic aspects of the self (e.g. thoughts, emotion and behavior) that allow for learning experiences (Tourangeau, Nord, Lê, Sorongon, Hagedorn, Daly, & Najarian, 2017; Carlson, Zelazo, & Faja, 2013; Diamond, 2013; Zelazo, 2015). Because EF is a series of critical cognitive functions/processes that allow for mathematics achievement, it will therefore be an important factor to consider to better understand the relationship between mathematics achievement and the role of demographic characteristics.

EF has a long history of being considered important for the basic skills (e.g. attention) to perform academic activities such as mathematics successfully. EF processes are considered central to learning and academic performance. Any shortcomings in these functions will likely stem from demographic characteristics or, in more severe cases, a diagnosis of a learning disorder in an academic subject (Zelazo, 2015). A reading disorder, for example, is manifested as a deficit in one of the processes that compose EF (Swanson, Mink, & Bocian, 1999). Similarly, mathematics learning difficulties have unique deficits that occur in the processes related to EF (Geary, 2013; Swanson & Alloway, 2012). Low achievers in mathematics have also been found to have shortcomings in different processes related to EF (Geary, 2011; Swanson & Alloway, 2012). These conclusions make EF a reliable indicator of effective learning, performance, and achievement. Because of the complexity of its functions, EF is represented by three distinct and unique mental processes that work independently to control/inhibit, hold in memory space, or reiterate, depending on the task, an individual's cognitive functioning (Zelazo, 2015). The names of these EF's are working memory, inhibitory control, and cognitive flexibility (Miyake, Friedman, Emerson, 2000; Zelazo, Muller, Frye, & Marcovitch, 2003). Each function is responsible for memory capacity, attention, and solutions to novel problems, respectively. The following will be a brief discussion of EF's functions, how these processes relate to mathematics performance, and how they might be influenced by demographic characteristics during early development.

This study broadens the focus of previous investigations that considered the effects of cognition (e.g. executive function) on mathematics achievement. An emphasis will be given to the influence of demographic variables that highlight the variability between groups. Because grade level is an indicator of critical periods for learning mathematics (Clark, Pritchard, & Woodward, 2010; Monette, Bigras, Guay, 2011), special attention will be given to the possible associations and changes over time between executive function and mathematics achievement. Of critical importance here is the overall EF-math relationship over the early childhood education period and the strength of the relationship when accounting for demographic characteristics (e.g. gender, race/ethnicity, SES, and grade level). An additional exploratory question that will be considered in this investigation is whether the EF-math relationship may vary because of demographic characteristics.

Demographic Differences in Mathematics Achievement and Executive Function

Demographic characteristics provide a source of variance that makes for unique experiences in all aspects of everyday life. In relation to academic achievement, these characteristics allow researchers to understand the differences between individuals and performance groups and address the setbacks or difficulties in underperforming groups. One example is the Monette et al., (2011) investigation that replicated the relationship between academic achievement and executive function while accounting for confounding variables. Some of those variables were socio-economic status, gender, ethnicity, and school grade. These are variables that have been found to impact mathematics achievement in different ways. Socio-economic status is important because it influences the degree to which families can provide the necessary resources and experiences to ensure the typical development of the child at different grade levels.

The gender disparity in mathematics at a later age and in the STEM fields overall is a significant and quite complex phenomenon (Halpern, Benbow, Geary, Gur, Hyde, & Gernsbacher, 2007). Also, it is worth noting that studies have not found differences when looking at the relationship between gender and overall EF for early grade levels (Monette, Bigras, & Guay, 2011). Some researchers have found ethnic differences in mathematics performance, while others attribute the differences to characteristics that carry more weight (e.g., SES). Thus, ethnicity will be included in this study. Because early childhood education is a critical period of development for many basic skills (Diamond, 2013; Zelazo, 2015), the fourth characteristic that will be considered here is grade level. Because so much of EF and mathematics achievement depend on the child's development, accounting for grade level will likely reveal the associations between these variables across the data time points. By taking this approach, the importance of demographic characteristics is addressed, and methodological precautions are taken to account for context between the relationship of academic performance and executive function. The following chapter will give a more detailed account on the importance of the demographic variables described by the literature.

Concern for Mathematics Achievement

Mathematics achievement has been a concern for many populations beyond researchers, teachers, and policymakers. Bundled with STEM education, mathematics achievement is regularly discussed in the media in several contexts from No Child Left Behind (NCLB) to its importance in basic academics. Internationally, the United States is regularly compared in mathematics performance and achievement to other developed nations (Hyde & Mertz, 2009). Given the importance of mathematics achievement and the implication for career options and SES (Morgan, Li, Farkas, & Cook, 2017), researchers have been studying this area for the better part of the late 20^{th} century. Indeed, all aspects of mathematics achievement have been under scrutiny. Anything from teaching mathematics, student mathematics performance, and mathematics achievement have been compared to other developed nations. This is one of the reasons researchers, educational institutions, and government agencies spend resources to better understand mathematics in academic institutions (Ashcraft & Krause, 2007). It is also important to note that, in all levels of education, mathematics continues to be prioritized in the United States (National Mathematics Advisory Panel, 2008). At the individual level, learning difficulties affect children at an early age (Geary, 2013; Swanson & Alloway, 2012). These conditions influence the trajectory of an individual's academic experience and even career choice. Given the importance of mathematics at different levels of society and at the individual level, it is more important to use the resources made available by technological advancements that improve on sample size, longitudinal information, methodological, and quantitative approaches to better inform learning at an early age.

Purpose of Investigating Factors that Influence Mathematics Achievement

The purpose of this study is to contribute to the literature by building on previous investigations, closely considering demographic characteristics (Little, 2017; Monette, Bigras, & Gray, 2011), and focusing on the relationship between executive functioning and mathematics achievement over a five-year period (Morgan, Farkas, & Wu, 2009). Like other work, this investigation will be framed in the context of comparing the longitudinal learning models that explain differences in academic performance (Morgan, et al., 2009). While substantial work has been done in this area, this remains an arguably open issue due to the difficulty of acquiring a large, representative sample, inaccessibility of continuous data from early academic years, and methodology to account for the nested nature of educational institutions. This investigation aims to better understand potential discrepancies found in other studies and contribute to the literature by replicating previous findings on the relationship between EF and mathematics achievement. Another purpose of this investigation is to highlight the normative development between EF and mathematics performance as influenced by demographic characteristics. In the interest of conducting a methodologically sound investigation, demographic characteristics will be used to better understand the sources of variability between individuals and performance groups. Of interest here are gender, ethnicity/race, and socio-economic status. These demographic characteristics will be carefully investigated for influence over mathematics achievement throughout early childhood education. This investigation will closely consider factors that have, at times, been dismissed, or have not been considered longitudinally. In an effort to expand on previous literature closely and explore the growth trajectory of mathematics achievement, this investigation will be guided by mathematics skill growth models (Aunola, Leskinen, Lerkkanen, & Nurmi, 2004). Using the ECLES-K:2011 Child Session investigation data will facilitate answers to research questions through the benefit of a large sample size and multiple waves of data collected with information that was not previously accessible to education researchers. This investigation will allow researchers, educators, and policymakers alike to better understand growth learning patterns and use that information to address potential mathematic achievement deficits during the critical period of early childhood education.

Benefits for Better Understanding Mathematics Achievement

The implications for this investigation are at different levels of analysis. The first benefit of this investigation is that it combines methodology from recent education research that has informed mathematics achievement. With the use of a large sample of nationally representative students from the United States over a five-year period, the dataset used here provides a wealth of information not available to social scientists and education researchers until recently. Another benefit to this investigation, when compared to other longitudinal inquiries, is that this investigation looks at the normative growth patterns at the beginning of formal education when basic skills are being developed (Aunola, Leskinen, Lerkkanen, & Nurmi, 2004). Secondly, access to cognitive performance variables, mathematics achievement outcomes, and demographic characteristics over time will give better insight into the relationship among these variables. Thus, the combination of access to growth data and variables that account for mathematics achievement allow for a more thorough investigation of the influences of mathematics achievement. This investigation aims to contribute to the understanding of mathematics achievement during the critical years of early childhood education.

Research Questions to Better Understand Mathematics Achievement

A series of research questions were developed based on the previous literature and theory reviewed. These research questions aim to improve our understanding of mathematics achievement over the early childhood education period.

1. *What is the average math achievement growth trajectory from kindergarten through fourth grade, accounting for the nesting of students within schools, and summer loss?*

2. *To what degree do gender, racial, or socioeconomic status differences account for variation in the intercepts, slopes, and summer drops in the mathematics trajectories at the student and school levels?*

3. *Is executive functioning associated with change in math achievement over this period?*

4. *Which student sociodemographic characteristics are most strongly associated change in math achievement, and to what degree does controlling for student sociodemographic characteristics alter the association between executive functioning and change in math achievement?*

5. *To what degree does math achievement growth vary among schools, controlling for demographics?*

Overall, this study aims to expand upon previous literature by further examining the relationship between executive functioning and mathematics performance, as well as the role of several key demographic variables. While much work has been done in this area, this remains an open issue as the findings have varied from study to study. This study hopes to expand upon the existing literature by exploring more complex relationships between demographics, executive functioning, and mathematics performance. We know there is evidence for racial, gender, and SES differences in mathematics performance, and that EF is a critical predictor of academic performance, but seldom have these factors been considered in the same investigation. This investigation will study the normal growth in math learning and EF trajectories in grade levels K-

5. Other investigations have studied similar age groups with a focus on learning difficulties (Morgan, Farkas, & Wu, 2009; Morgan, Li, Farkas, Cook, Pun, & Hillemeier, 2017). However, this investigation will focus on the typical development of cognition and mathematics achievement in children grades K-4. Another contribution is that no other investigation has studied the 5-year growth trajectories of typical growth in mathematics achievement and EF. This dissertation will address that gap in the literature. The following chapter will discuss in greater detail the previous literature and theory relevant to this investigation's research questions and hypothesis.

CHAPTER 2

LITERATURE REVIEW

In this chapter, there will be a thorough literature review on the importance of mathematics achievement and the factors that contribute to success in that domain. First, there will be a discussion on the reasons why children perform poorly in mathematics. Following will be a discussion on the cumulative growth model and lag models. These theoretical models provide different perspectives for understanding growth in mathematics achievement and performance and provide context for understanding the achievement gap. Factors to consider in the literature when describing mathematics achievement are gender differences, racial differences, and socioeconomic differences. These demographic characteristics provide the context for understanding differences in academic performance and mathematics achievement. Because this investigation seeks to explore as many factors available to understand mathematics performance, executive functioning and its components will also be thoroughly reviewed. The role of the cognitive domain in learning mathematics is critical to understanding overall mathematics performance. The chapter will conclude with a summary of the literature that describes how the different factors that affect mathematics performance interact and how time affects these interactions.

Assumptions Concerning Mathematics Performance and Low Achievers

First, it is important to clarify some assumptions about mathematics performance. Under the umbrella of STEM, mathematics and science achievement are among the academic topics that frequently make the news, with the general conclusion being that the United States underperforms (Ashcraft & Krause, 2007). Similarly, student academic performance is relatively low in the United States compared to many of its counterparts. In addition to this, there is the necessity of experts in the fields of education and social science, as in any technological society, to better understand academic achievement. All of these reasons highlight the importance of examining mathematics performance, and specifically attempting to zero in on the reasons behind poor performance in mathematics among students in the United States (Ashcraft & Krause, 2007). Some research suggests children perform poorly due to learning disabilities that interfere with their ability to learn math. Other explanations suggest that poor performance is due to economic disadvantages that do not allow for the opportunities for proper academic development. It is therefore even more important to investigate mathematics achievement during the critical period of early education.

For example, it has been found that kindergartners that underperform in mathematics are less likely to attend college as adults, own their own home, be employed, or live in higher income neighborhoods. The same research also suggests that exhibiting poor math skills increases the likelihood of lower wages and lowers employability for adult workers more than other factors, such as low IQ and poor reading skills (Morgan, Li, Farkas, Cook, Pun, Hillemeier, 2017). Research also suggests that cognitive factors also play a role, such as a child's EF skills, where cognitive flexibility and working memory predictably account for educational achievement later in life. Therefore, a child's EF is a strong predictor of academic achievement. This investigation aims to assess whether a child's cognitive flexibility and working memory impact academic achievement in mathematics, and their subsequent growth, by accounting for as many factors as possible. The following sections of this chapter explore current and past research to determine whether measures of working memory and cognitive flexibility can be used, and to what extent these measures can predict a child's math achievement throughout their education.

Achievement in Mathematics

Mathematics achievement can be attributed to several factors related to cognition and demographic characteristics. How, and to what degree, these factors contribute to the achievement of mathematics is not clear. The literature offers contrasting views on the process by which children become proficient in mathematics. Identifying proficiency is clear for typical achieving and high achieving groups. However, individuals with difficulty or a more severe disability in mathematics are more challenging to identify (Geary, 2013; Swanson & Alloway, 2012). One approach has been the use of discrepancy formulas and other exclusionary criteria that have traditionally been utilized to differentiate between individuals who display math difficulties and those who underperform in mathematics as a result of demographics-related disadvantages. This realization excludes the notion that mathematics difficulties, overall, might be due to low IQ (Geary, 2013; Swanson & Alloway, 2012). Therefore, it is the case that math difficulties occur in individuals who perform poorly in math achievement assessments relative to the IQ of these same individuals. On the other hand, using a discrepancy formula to identify learning disabilities is considered increasingly untenable (Morgan, Farkas & Wu, 2009). Researchers have concluded that identifying mathematics achievement requires an array of skills related to the cognitive processes that allow individual differences in performance is less understood (Clark, Pritchard & Woodward, 2010).

Math proficiency allows an individual to actively reason problem elements in order to arrive at a possible solution. When solving a problem, an individual must represent information as part of the working memory process and shift attention adequately to problem elements (e.g., the different components of a problem). On the other hand, there are aspects of mathematics ability that are dependent on factual knowledge that can be readily retrieved from an individual's long-term memory. Mathematics is an inherently effortful academic domain that makes explicit demands on EF abilities of very young children during a critical part of academic development (Blair & Razza, 2007).

Because the skills being developed are critical, young children who experience learning difficulties in math are at risk for having fewer societal and educational opportunities throughout subsequent years. Kindergartners who perform poorly on math assessments are less likely to attend college, own their own home, be employed, or live in higher-class neighborhoods. Learning difficulties in math are also associated with a child's increased risk for feelings of isolation, inferiority, and socio-emotional maladjustments (Morgan, et al., 2017). The inability to be proficient in math is a large obstacle when it comes to societal opportunity. For instance, having poor math skills is associated with lower income earnings and employability more than any other factor, such as low IQ or poor reading skills. Typically, a child is believed to have math difficulties if their performance is between the 25^{th} -30^{th} percentile on mathematics assessments (Morgan, Farkas & Wu, 2009, p. 307).

Executive Function

Executive functioning refers to the interdependent processes that orchestrate the regulation of emotion, behaviors, and cognitions (Zelazo, 2015). Zelazo, Muller, Frye, and Marcovitch (2003) view EF as a functional construct associated with the psychological processes required for task-oriented demands (p. 2). According to Monette, Bigras, and Guay (2011), EF is a set of processes that typically include working memory, planning, flexibility, concept formation, and fluency. EF skills are neurocognitive skills that require higher cognitive processes and task-oriented focus (Zelazo, 2015). Furthermore, Clark, Pritchard, and Woodward (2010) define EF as a series of processes required for complex problem solving and task-oriented behavior (p. 1177). For the purpose of this investigation, EF refers to cognitions that allow for the required mental strain of problem solving for children to successfully solve mathematics problems.

EF is a strong predictor of academic achievement. Previous findings have shown that preschool is a critical period for the development of EF, specifically around the age of four when the organization of the cognitions critical to EF reportedly takes place (Monette, Bigras & Guay, 2011, p. 159). Other findings report that the cognitions in preschool predict later scholastic achievement (Monette, Bigras & Guay, 2011). Little (2017) found that EF skills also help an individual focus on goal attainment, a critical component of academic success. These skills are commonly referred to as the "air traffic control system of the brain" (p. 103), where instincts and thoughts are processed and sorted throughout the brain. The processing that occurs is a result of three different skills that make up EF: working memory, cognitive flexibility, and inhibitory control (Zelazo, 2015). Several researchers have posited these skills to be the main components of EF (Monette, Bigras & Guay, 2011). Working memory and cognitive flexibility are the processes that will be critical for this investigation.

Executive Function Skills

The components of EF, that is, working memory and cognitive flexibility, work together to facilitate problem-solving. These skills play a critical role when making inferences, characterizing problems, and keeping plans and goals in mind (Zelazo, 2015). The manifestation of strong EF is observed in young children with the skills to maintain attention, plan better, remember and apply instructions, and multitask. Deficits in these skills result in difficulties with learning because they interfere with a child's classroom success. For example, both cognitive flexibility and working memory deficits have been found to interfere with a child's comprehension monitoring, where a child may struggle to integrate and hold information in their long-term memories while also attempting to make and process inferences. The literature has produced substantial evidence to suggest that children who exhibit learning difficulties have poorly developed EF skills (Morgan et al., 2017).

Working Memory

Tasks that rely on working memory have been characterized as maintaining mental information while simultaneously working on a task (Colom, Flores-Mendoza, Quiroga, & Privado, 2005). This skill or function is distinct from tasks that only require the use of short-term memory, as tasks involving the use of just short-term memory only necessitate maintaining information without having to process it. The role of working memory on mathematics specifically as well as cognitive functioning more generally, cannot be understated. Within the resource sharing model, a trade-off exists between storage and processing (Daneman & Carpenter, 1980). While focusing on reading, Daneman and Carpenter posit that differences in comprehension may be due to individual variation in working memory capacity, and created a test which places large stresses upon both processing and storage in order to further explore this trade-off. They found that the scores on their listening span task correlated with other reading comprehension measures that had correlations of similar strength (Daneman & Carpenter, 1980).

The importance of working memory extends beyond task completion. Previous research has reported a relationship between short-term memory, working memory, and general intelligence. One such study conducted by Colom, Flores-Mendoza, Quiroga, and Privado (2005) found a strong association between short-term memory and working memory. Working memory was found to be a strong predictor of general intelligence compared to short-term memory. It was also found that the relationship between working memory and general intelligence is attenuated by short-term memory, suggesting that the short-term memory component of working memory moderates this relationship. Within this study, working memory was measured using the ABCD and the Alphabet tasks, which were modeled after the CAM battery. Additionally, quantitative working memory was measured using the Mental Counters and Computation Span tasks. Spatial working memory was measured using the dot matrix and letter rotation tasks. A similar study served to further explore the relationship between short-term memory and working memory (Colom, Shih, Flores-Mendoza, & Quiroga, 2006). Here, the authors continued to examine the extent to which working memory and short-term memory overlap by using a set of 12 memory span tasks, with half of these tasks serving to measure short-term memory, and with the other half measuring working memory. Within this study, verbal working memory was measured using the ABCD and alphabet tasks, with quantitative working memory measured using the mental counters and computation span tasks. Spatial working memory was measured using the dot matrix and letter rotation tasks. The results of this study found that short-term memory and working memory largely overlapped, and that the limitations underlying these domains were largely shared (Colom, Shih, Flores-Mendoza, & Quiroga, 2006).

Given the importance of working memory in information processing, it is therefore important to consider this domain a critical factor for academics and mathematics performance. Indeed, according to Morgan, et al., (2017), deficits in working memory, or information processing, may impede a child's academic performance. This is due to the fact that mathematics tasks require children to manipulate and store information. On the other hand, researchers found that deficits in cognitive flexibility may not be as important during the early years in determining academic performance. For example, during the early years of education, few classroom tasks require multi-step problem solving and thus place fewer demands on cognitive flexibility deficits. The next section will discuss the current literature on the EF component of cognitive flexibility.

Cognitive Flexibility

Cognitive flexibility refers to thinking about a task in several ways (Zelazo, 2015). According to Little (2017), cognitive flexibility is an individual's ability to switch their attention or perspectives. It is an important feature of EF that allows for different perspectives towards daily demands (Handbook of Behavioral Neuroscience, 2016). Another way of interpreting this function, according to Diamond (2013), is that cognitive flexibility is the ability to think outside the box and see one item or aspect from a different perspective. Therefore, this function refers to an individual's flexibility and quickness at adapting to changing circumstances (Diamond, 2013).

Cognitive flexibility has been defined as "the readiness with which the person's concept system changes selectively in response to appropriate environmental stimuli" (Scott, 1962: 405). It has also been defined as "the ability to shift between response sets, learn from mistakes, devise alternative strategies, divided attention, and process multiple sources of information concurrently" (Anderson, 2002: 74). Cognitive flexibility relates to the ability of an individual to change their working strategies based on a change in task demands (Singer, Ellerton, & Cai, 2015). This skill is composed of three main constructs: cognitive variety, cognitive novelty, and change in cognitive framing. Cognitive variety relates to the amount of diversity in mental templates for problem-solving, which are associated with cognitive pathways or perspectives. Cognitive novelty relates to an individual's focus during study and content mastery and the incorporation of additional external perspectives. Cognitive framing relates to persistence when attempting to solve a new problem using an old method (Singer, Ellerton, & Cai, 2015).

Cognitive flexibility is also an important indicator of development as it improves over time for individuals. While infants frequently exhibit perseverative behavior, this declines in childhood, becoming rare in adolescence. The ability to rapidly switch between different response sets is first seen among children aged approximately three or four years, while children of this age still have trouble when faced with rules of increased complexity. Even seven-year-old children have difficulty switching behavior when there are multiple dimensions, though this limitation has been found to improve between seven and nine years of age. Cognitive flexibility then continues to improve during adolescence. The ability to learn from mistakes and produce different strategies to solve a problem is first seen in early childhood and continues to develop as children age (Anderson, 2002).

Factors Influencing the Development of Executive Function

EF skills develop within an environmental context, especially throughout early childhood. Brain development can be impaired by toxic stress, ultimately impacting the development of the skills that make up EF. On the other hand, stimulating and rich environments provide the necessary experiences for proper development of EF (Little, 2017). Established structures and routines in a child's life, for example, help promote EF skills. However, not all children are raised in an environment that promotes and encourages routines and structures, making the environment an important factor in academic achievement. Later in the chapter will be a discussion of the environmental factors that have been deemed important for the proper investigation of the relationship between mathematics achievement and the executive function.

The Relationship Between Executive Function and Mathematics Achievement

The literature has reported a direct association between a child's mathematical skills and their EF skills. For example, studies show that inhibitory skills in preschool can predict a child's mathematical ability (Bull, Espy & Wiebe, 2008). Furthermore, studies have shown a distinct association between executive function skills and concurrent math performance (Clark, Pritchard & Woodward, 2010). Research also suggests an association between academic achievement and EF. Math skills, however, are most closely linked to working memory (Monette, Bigras & Guay, 2011). Morgan et al., (2017) found that deficits in both cognitive flexibility and working memory in kindergarten increase the risk of math difficulties in the first grade. Worth pointing out is the strong association between the risk of deficits in working memory. Thus, these researchers suggests that multi-component interventions may help young children with math difficulties, especially with deficits in working memory.

Blair and Razza (2007), for example, found a moderate to strong relationship between mathematical ability and EF. This finding is consistent with the neuroscientific literature suggesting that there is an association between neural substrates for the functions of EF, numerical ability, and quantitative reasoning (p. 658). Furthermore, on a sociological level, the relationship between mathematical ability and EF are consistent with the historical evidence which indicates a shift in mathematics curriculum during the early elementary grades. Historical reviews of United States public school textbooks show a pattern of visual-spatial working memory and pattern completion problems that demand the use of EF to complete different components of a problem (p. 658).

Growth Trajectories

Previous research has focused specifically on changes in development over time. One study by Morgan and colleagues (2009), using data from the ECLS-K, examined the relationship between mathematics difficulties in kindergarten and growth trajectories from the first through fifth grade. This investigation found an important association between these metrics. Specifically, persistent mathematics difficulties were associated with the lowest growth rates in later grades. Children exhibiting mathematics difficulties in either the spring only, or the fall only in their kindergarten year, had the second and third lowest growth rates, respectively. This suggests that the capacity and performance of mathematics difficulties in kindergarten is a predictor in later mathematics proficiency (Morgan, Farkas, & Wu, 2009). The study also found changes in growth over time based on mathematics difficulties, suggesting an interaction between mathematics difficulties and time on mathematics performance. Specifically, the greatest degree of growth was found over time among children who did not exhibit mathematics difficulties at either time point, with the lowest growth over time evidenced among those exhibiting mathematics difficulties in both the spring and fall. They suggest the importance of early interventions among children exhibiting mathematics difficulties to help increase their mathematics performance as elementary school students. A child's developmental trajectory is defined by interactions between development, exposure to learning, experiences, and mental development (McClelland, Cameron, Duncan & Bowles, 2014, p. 3).

Monette, Bigras and Guay (2011) attempted to determine if EF in young children was associated with first-grade school achievement. EF measures were administered to 85 kindergartners, 46 girls and 39 boys aged 5-6 years old. School achievement was measured at the end of first grade. The researchers reported that a child's first-grade reading/writing and math skills were associated with kindergarten EF. However, only working memory was reported to contribute to overall school achievement. The researchers believe the relationship between math achievement and working memory is determined by the problems in mathematics. For instance, children typically handle math problems using mental models (i.e., visuospatial representations) instead of symbolic language; skills that develop as the individual matures and continues throughout the critical period of early childhood education.

While several studies have shown an association between young children's inhibitory skills and mathematics achievement, the relationship is less clear in older children. The current consensus in the literature suggests a close link between working memory and mathematics achievement. It is unclear, however, whether these abilities contribute to the identification and prediction of later mathematical skills. Bull, Espy and Wiebe (2008) examined whether measures of working memory and short-term memory in preschool children could account for future academic achievement. Children were assessed using cognitive measures and reading and mathematics assessment outcomes. Researchers found that the visual-spatial and short-term memory span predicted math ability at 7 years of age. The EF skills were found to predict learning in general, but not a specific domain. EF, as an overall process, can be used as a predictor of a child's future academic achievement.

Theoretical Models for Understanding Growth in Mathematics Achievement

The current literature proposes two theoretical models that describe how students gain their mathematical knowledge and the resulting mathematical achievement. These models propose that students gain knowledge, but due to demographic characteristics and outcomes in achievement, result in different performing groups. The cumulative growth model predicts a continued achievement gap. Alternatively, the lag model expects that students who underperform early on will have rapid growth and catch up to their higher-achieving peers. The following is a description of the two models that will be implicitly tested in this investigation.

Cumulative Growth Model

The cumulative growth model explains how children gain their understanding of mathematics. Throughout their ongoing interactions with parents, siblings, teachers, peers, and others, children continually extend and refine their early mathematical understandings. Their learning is typically completed through information instruction, such as helping a child learn how to count, educational material such as songs, etc., which are provided to children during preschool years. Research has shown that children who enter kindergarten with mathematical knowledge continue to increase their knowledge, while children with less mathematical knowledge learn mathematical concepts at a slower rate. One clear manifestation of this model shows that children who experience an early onset of math difficulties will continue to display math difficulties as they progress through elementary school (Morgan et al., 2009).

Lag Model

According to the lag model, children who enter kindergarten with less educational training are more likely to increase their mathematical skills and knowledge more rapidly than children entering kindergarten with higher levels of math knowledge. Consequently, children with lower levels of math skills begin to catch up to their peers with higher level math skills. This process occurs when children with lower math skills start to receive systematic instruction, which helps children who have experienced learning difficulties prior to entering kindergarten. Over time, the magnitude of the gap between higher and lower skilled children should decrease instead of increasing or remaining constant. More severe cases of children with early onset of math difficulties may not even display symptoms when they are older because the deficit has been addressed. Thus, according to the lag model, deficiencies during early education—if properly addressed—may not predict academic achievement throughout the child's education. Several studies have supported the lag model, indicating some children who enter kindergarten with low levels of math skills display moderate to rapid growth in math skills as they get older (Morgan, Farkas & Wu, 2009). While this is an accepted model, the lag model does make it difficult to understand how math skills prior to entering school impact later education.

The models discussed above are theoretical descriptions of growth in mathematics achievement and performance. Evidence supporting both theoretical models has been found in the literature (see Morgan, Farkas & Wu, 2011). This investigation will use the models to guide the research questions to better understand instances where the cumulative growth model and lag model apply. Understanding how these models can be applied in the context of this investigation, and under what circumstance, will inform the best way to target student learning in mathematics and academic performance. The theoretical models discussed also make the case that learning experiences outside of school, such asin the home or through social interactions, are key to learning and therefore mathematics achievement. The following sections will discuss the importance of accounting for gender differences, racial differences, and socio-economic differences.

Gender Differences

The complexity of gender as a demographic characteristic has confounded researchers studying academic achievement. Indeed, researchers have written articles on gender in relation to mathematics performance and EF where the findings have not been clear and sometimes conflicting. In one such paper, Halpern et al., (2007) reviewed the literature to answer the tough question of how gender differences affect performance in science and mathematics. The researchers reported some insights into potential gender differences, though they emphasized that the report should be interpreted with caution. The authors noted that, developmentally speaking, infants should not show differences in EF because they have not had many experiences. However, with a large enough sample, they found that female infants performed better on EF tasks. At later grade levels, gender differences have been reported in the literature when looking at the performance of EF (Bull et al., 2008). In elementary school, female participants again outperformed male participants in reading tasks. In relation to the visuospatial sketchpad, male students outperformed female students as early as preschool (Levine, Huttenlocher, Taylor, & Langrock, 1999). Halpern et al., (2007) also reported that while differences are observed in EF at an early age, and while male students tend to mostly represent the higher performing groups in preschool mathematics, there is no clear evidence that males are better in overall mathematics performance. Despite the perceived disadvantage, female elementary school students outperform male students in mathematics. The researchers did note a pattern, where female students are expected to outperform male students in specific mathematics tasks related to computational knowledge and speed during the early grades of elementary school. That the differences in gender due to mathematics performance and EF are not clear cut reflects the complexity of gender in attempting to explain achievement differences. Including gender as a covariate and as a moderator will provide insights into the potential differences and how these differences are likely to vary across time in relation to EF and mathematics achievement.

Other research has also indicated gender differences in mathematics and science achievement, with boys having significantly higher achievement compared to girls. Perhaps most notably, an expansive meta-analysis conducted using data from the National Assessment of Educational Progress focused on gender differences in mathematics and science achievement scores for American students during the 1990-2011 period (Reilly, Neumann, & Andrews, 2015). Achieving a total sample size of nearly three million students, the results of this study found small but reliable gender differences in mathematics and science achievement scores, with females having significantly lower scores than males. The researchers also investigated whether gender differences were similarly present in high achievers in these areas of study (Reilly, Neumann, & Andrews, 2015). Gender differences were found to not significantly change based on grade level with respect to science, though gender differences were found to significantly increase depending on grade level with respect to mathematics achievement scores. Additionally, there was no significant effect of year, nor was there an interaction between year and grade level, indicating that gender differences were relatively consistent across time (Reilly, Neumann, & Andrews, 2015). This investigation contradicts the Halpern et al. (2007) review, though it is consistent with the Levine et al. (1999) findings.

Other research has found substantial changes in gender differences over time on performance. Focusing on science test scores more generally, research examining changes over the course of childhood education illustrate growing effect sizes between males and females. For example, effect sizes have been found to be the smallest in the cases of youngest students, with 5th graders having an effect size of $d = 0.10$ (Maerten-Rivera, Myers, Lee, & Penfield, 2010), with this increasing to $d = 0.26$ in the case of 10th grade students (Mau & Lynn, 2000), and increasing further to *d* = 0.28 with respect to 12th graders (Mau & Lynn, 2000). These findings suggest that age moderates the effect of gender on test scores. It is therefore important to include gender in the model during the critical years of development (k-4th grade) to explores the potential differences in gender performance.

To explain differences in mathematics achievement, it is important to consider factors affecting academic performance. For example, mathematics anxiety is defined as the self-imposed pressure that influences the processing of information (e.g., numbers), therefore affecting the overall solution of a math problem in academic contexts (Passolunghi et al., 2016). Other investigations have also reported gender differences in math performance and anxiety. These findings reported that males are more likely than females to choose majors that lead to STEM careers (Halpern et al., 2007). It has also been suggested that individuals with math anxiety will be more likely to have negative attitudes about their capabilities at solving mathematics problems (Lent et al., 1991). These negative attitudes may then serve to increase math anxiety and math avoidance (Passolunghi et al., 2016). It has also been suggested that poorer math performance among females has the effect of reducing their confidence and interest in math and science, which will negatively impact their career choices by pushing them away from these areas (Ganley, Vasilyeva, & Dulaney, 2014). These findings are important because research has also found math anxiety to be associated with reduced short-term memory as well as poorer working memory (Passolunghi et al., 2016). Overall, a large body of literature exists suggests that anxiety serves to reduce the capabilities of working memory (Peters, 2015). Stereotype threat has also been suggested as a possible reason behind the gender differences seen in mathematics performance (Ganley et al., 2013). Stereotype threat relates to individuals being impacted by an unconscious fear of behaving in a way which would serve to confirm a negative stereotype relating to their performance in some specific domain. With respect to mathematics, this would consist of the stereotype that males perform better than females. However, studies that have tested for the potential of stereotype threat as being a factor in the differences in mathematics performance based on gender have not yet revealed any evidence for this being a factor. This may be since stereotype threat is only relevant in very specific circumstances, or that it acts as a pervasive, all-encompassing factor (Ganley et al., 2013).

Consistent with previous findings, it is expected that gender differences will be observed in this investigation. It is further expected that male students will outperform female students in overall mathematics. If differences are observed in other assessments related to EF, the findings will potentially reveal more information about gender performance.

Racial Differences

The literature has reported contradictory findings on racial differences in mathematics achievement and EF performance. For example, Little (2017) found that racial gaps in cognitive flexibility and working memory were narrower than the gaps found in academic achievement. This was especially true for Hispanic students who were between kindergarten and second grade. For instance, the gap between White students and Hispanic students on the Numbers Reversed task dropped from 0.59 to 0.27 standard deviations (Little, 2017, p. 104). In mathematics performance, the gap between White students and Hispanic students was marginally narrow, from 0.64 to 0.54 standard deviations (Little, 2017). Thus, according to this study, there are racial differences in EF skills reported in the literature for young children. In another investigation, Hooper, Roberts, and Sideris (2010) examined the social-behavioral predictors of math skills and math trajectories between Caucasian and African American students. The results of the study reflect the contribution of early math skills to later academic performance in both Caucasians and African Americans. Thus, the findings from this study suggest that there is not a racial difference when it comes to academic achievement in math.

In another investigation, however, Sonnenschein and Sun (2017) assessed the relationship between a child's early math skills and their parent's knowledge of the child's development. The study used the Early Childhood Longitudinal Study-Birth Cohort in order to assess the differences between Asian, Latino, Black and White children's early math skills at the entry of kindergarten. In addition, the researchers wanted to investigate if the parental understanding of development and experiences at home accounted for the relationship between race and mathematics skills (Sonnenschien & Sun, 2017). Knowledge of the child's development by the parents was assessed when the child was 9 months old. The experiences at home, such as home enrichment and literacy, were assessed during preschool. White and Asian children entered kindergarten with higher math scores than Latino or Black children. There were also significant differences regarding the frequency of engaging in home enrichment and literacy activities. Thus, the association between math scores and race/ethnicity was mediated by their parents' knowledge of their home literacy activities and development. Researchers stress the importance of educating parents on educationally relevant activities and how to foster a child's math skills by engaging in such activities to close the racial/ethnic gap in mathematics.

Not unlike previous findings, it is expected that racial differences will be observed on the EF tasks and manifest in the mathematics achievement outcome measures. Because development is critical for performance, the differences between the groups are expected to be observed as early as kindergarten and continue throughout the investigation.

Socioeconomic Differences

The literature has pointed out that a child's risk of developing math difficulties is likely due to their standard of living manifested in a lack of opportunities to learn. According to Morgan, Farkas, and Wu (2009), for example, children from low-income households are more likely than children from high-income households to display poor math skills. On the other hand, some studies have yielded contradictory results. One study reported that children from low-income areas were not more likely to exhibit math difficulties than children from higher income areas. Researchers suggest the contradictory findings are due to methodological limitations. For example, only a few studies have investigated the direct impact that a child's socioeconomic status has on mathematical learning without using the reported income or education of the parent or guardian. Instead, the impact of socioeconomic status has been measured indirectly through information gathered from the academic institution the child attends and their eligibility for a free lunch.

The racial differences in mathematics performance can be explain by considering differences in EF because of SES. Children from lower SES generally perform lower on EF tasks (Zelazo, 2015). Little (2017) explored the socioeconomic and racial differences in young children regarding EF. The researcher used the Numbers Reversed task and the Dimensional Change Card Sort (DCCS) to measure working memory and cognitive flexibility, respectively. Performance differences were reported in cognitive flexibility and working memory skills during the first year of kindergarten because of SES. The results of the study showed that students in the top quintile for socioeconomic status scored higher than the students in the lowest quintile on the Numbers Reversed task. It is important to note that the same result was found for race. While these gaps did decrease over time, the decrease in working memory was much more pronounced than cognitive flexibility. Furthermore, the socioeconomic status gaps between measures of academic achievement and measures of working memory are similar at kindergarten entry. That being said, performance differences in mathematics that are due to socioeconomic status are stable by the end of the second grade and decrease universally for working memory (Little, 2017).

Other Factors Influencing Mathematics Performance

There are other factors that can be associated with math achievement in young children. Other studies have identified risk factors associated with math difficulties including poor reading skills, entering kindergarten too young, and inattentiveness. These factors can also impact a child's math achievement. Furthermore, children are more likely to have mathematical difficulties if they repeatedly perform poorly on math knowledge measures instead of a single instance of underperformance (Morgan, Farkas & Wu, 2009, p. 308). While the summary on factors that influence mathematics achievement is not meant to be exhaustive, it shows that there are a myriad of factors that are important. Other factors are being explored and found by researchers to play a critical role in mathematics achievement. What is important here is to use the data that are available to best explain mathematics performance.

Applications and Implications

The importance of mathematics was highlighted earlier in this chapter. Mathematics and science are a mainstay of any technological society, and to stay competitive in the modern world, excelling in these areas is critical. However, American schools have not been found to excel in teaching either mathematics or science, with performance among American students being sub-par (Ashcraft & Krause, 2007). Ashcraft & Krause (2007) highlight the importance of working memory in everything relating to arithmetic and mathematics. They state that little research has been done examining math processing beyond that of basic arithmetic, which leads to a large gap in our knowledge relating to the relationship between EF and mathematics as it relates to anything more advanced than basic arithmetic problems (Ashcraft & Krause, 2007).

Conclusion

Math proficiency requires an individual to actively reason problem elements to arrive at a possible solution. Previous research has found a direct association between a child's mathematical skills and his or her EF skills. Throughout their ongoing interactions with parents, siblings, teachers, peers, and others, children continually extend and refine their early mathematical understandings. Mathematical skills and achievement, however, are not homogenous. Differences in mathematical skills and achievement are reflected when children enter kindergarten and have higher math achievement outcomes at the end of the first grade. On the other hand, research also suggests that children with lower math skills catch up to their higher-achieving peers when they start to receive systematic instructions.

As discussed, mathematical skills and mathematical achievement are most closely linked to the functions that make up working memory. The relationship between mathematical ability and EF are consistent with the historical evidence that indicates a shift in mathematics curriculum during the early elementary grades. In other words, the development that occurs in EF in early childhood is associated with the growth in learning and achievement of mathematics. One study reported that a child's first-grade reading/writing and math skills were directly associated with previous performance (e.g., kindergarten EF). However, of all the skills associated with EF, only working memory contributed to school achievement. Furthermore, researchers also found that children with high EF performance were more likely to perform well in math assessments at an early school age. Lastly, findings from several studies provide empirical support to address learning difficulties in young children through remediating the EF deficit, specifically working memory. Thus, while cognitive flexibility did impact a child's mathematics achievement, working memory was reported to have the greatest impact on academic achievement.

To account for environmental factors influencing academic performance, researchers have reported that poor performance can result from any economic disadvantages in the child's family that result in fewer learning experiences. Research has shown that environments that provide less opportunities for a child to learn mathematics can delay a child's math acquisition skills, which is remediated when the child enters school. However, there is a lack of consensus amongst researchers on whether math or EF skills are impacted by gender, race or socioeconomic status. Scholars have identified a gender gap in math skills in kindergarten and in first grade. These studies also suggest that a child's risk of developing math difficulties is due to their standard of living. Another contributing factor to the gender gap is from teachers favoring males during the earlier school years. Other studies have identified risk factors associated with math difficulties that include having poor reading skills, entering kindergarten too young, and inattentiveness. Regarding racial differences, research suggests that racial gaps in cognitive flexibility and working memory are narrower than the gaps found in academic achievement. It is worth pointing out, however, that the differences in cognitive flexibility and working memory skills at school entry are largely a result of SES.

Overall, deficits in working memory may impede a child's academic achievement. Also, children from lower SES typically exhibit lower levels of EF skills. Academic achievement is most closely associated with the working memory function of EF. Due to the level of complexity of preschool/kindergarten math, research has shown that a child does not need to exhibit complex cognitive functioning skills to find math solutions. Cognitive flexibility is required in later years, however, with more advanced math problems that require more thought processes. Thus, working memory has been found to be the best predictor of a child's later educational attainment.

CHAPTER 3

METHODS

Chapter 3 will discuss the theoretical background that has guided this investigation, detailed information about the data, assessments used to collect the data, and the methodological approach. An explanation for the data, variables of interest, and statistical approach used for the analysis will follow. The contrasting learning models used in this investigation in combination with the availability of the ECLS-K: 2011 dataset allowed for the unique opportunity to answer the proposed research questions. How the data was coded will be explained, from participant selection and missing data to selection of data time points. To address the hierarchical nature of the data, statistical equations describing the multilevel statistical analysis will be used to guide the reader through the analyses.

Conceptual Framework/ Theoretical Comparison of Mathematics Growth Learning Models

The theoretical framework that guides the research questions for this investigation are the cumulative growth model and the lag growth model used previously to learn about populations with learning difficulties (Morgan, Farkas, 2009), and described before in previous work (Aunola, Leskinen, Lerkkanen, & Nurmi, 2004). These models provide theoretical accounts that describe growth in mathematics learning achievement and performance.

The cumulative growth model hypothesizes that children gain insight into mathematics via experiences that result from demographic characteristics (e.g., gender, grade level, ethnicity, and SES). More specifically this model proposes that growth in mathematics learning results from social interactions with family members at home and with peers and instructors at school (Morgan & Farkas, 2009). This account is not unlike the zone-of-proximal-development described by Vygotsky (1987). The cumulative growth model, as the name suggests, predicts that the knowledge that one enters preschool with is built upon. Morgan and Farkas, (2009) make the case that those individuals with more knowledge are more likely to pick up mathematics skills faster and be more skilled and knowledgeable in mathematics. Similarly, individuals who start off preschool with less skill in mathematics will be behind those that are better prepared.

In contrast to the cumulative model is the lag model. The lag model is presented as another possibility that speaks to the benefits of the educational system, such as high-quality preschool programs (Welch, Nix, Blair, & Nelson, 2010). From this point of view, students start at different performance levels as a result of demographic characteristics. However, instead of the achievement gap remaining, the students who “lagged” behind gain performance ability and eventually perform at the same level as their peers. Morgan and Farkas (2011) point out that in the best-case scenario, students who at one point showed symptoms of a learning difficulty should perform average.

In a review of the literature, Morgan and Farkas (2011) point out that there is limited evidence for the cumulative performance model and the lag performance model. There is evidence in the literature to support the model that argues for the achievement gap, and the model that argues for overtime closure of the achievement gap. Previous investigations that studied growth trajectories in mathematics performance, however, have typically looked at two to three years of academic growth. This investigation aims to start the investigation at preschool and continue to the fourth grade to expand the test of mathematics performance growth models. The learning growth models will, therefore, guide the longitudinal questions, the expected outcomes, and statistical analysis approach.

Data Source / Participants

The Early Childhood Longitudinal Studies (ECLS) program includes three overall investigations. ECLS is sponsored by the National Center for Education Statistics (NCES), a program that is part of the United States Department of Education’s Institute of Education

Sciences (IES). The other studies in the Early Childhood Longitudinal Study (ECLS) program are the Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K) and the Early Childhood Longitudinal Study, Birth Cohort (ECLS-B) (Tourangeau et al. 2017). This investigation was made possible because of the most recent dataset that includes the early childhood longitudinal study, Kindergarten Class of 2010-2011 (ECLS-K:2011). The dataset is an across-the-board bank of information that contains wide-ranging topic areas from learning, development, school readiness and information about growth and change to cognition, social and physical development, experiences, and educational outcomes. All participants started when they were five years of age (e.g., 2010-2011 school year) when the children were in kindergarten. The sampling continued through the spring of 2016 when most of the participants were in the fifth grade (Tourangeau et al. 2017). The total sample size is 18,170 participants. Data were collected in the United States from 90 primary sampling units (counties or groups of counties) with an emphasis on ensuring that metropolitan and other areas, and both public and private schools were represented in the data collection. The total number of schools sampled was 1,310. A total of 23 students were randomly selected from each school. The sample included children from different racial/ethnic and socioeconomic backgrounds (Tourangeau et al. 2017). This included first-time Kindergarteners and repeating Kindergarten students. In addition, data were collected from sources other than children, such as parents/guardians, classroom teachers, special educators, school administrators, and, for Kindergarteners only, before- and after-school care providers. This investigation will focus on the child data.

Because the purpose of the assessment was to better understand skill and knowledge development, assessments occurred at multiple time points. Children participated in the investigation via direct assessment, which lasted approximately 60-80 minutes. From Kindergarten through second-grade, children were assessed for reading, math, executive function, science, and height and weight. When the participants reached the third grade, the assessment session was a child questionnaire. The first wave of data collection started in Fall 2010 and Spring 2011 for the Kindergarten year. The second wave occurred Fall 2011 and Spring 2012 for First grade. The third wave occurred Fall 2012 and Spring 2013 for second grade. The fourth wave occurred in Spring 2014 for third-grade. The fifth wave of data collection occurred in Spring 2015 for fourth grade. The final fifth-grade data collection wave occurred Spring 2016. All available time points will be included in the present analyses, which will account for the greater spacing of the annual repeated measures in later grades versus biannual repeated measures in earlier grades. Analyzing both the Spring and Fall data from earlier grades maximizes the information available early in development when more rapid changes may be occurring. The models described below will also allow for detection of a loss in gains over the summer when students do not receive formal instruction. All available subjects in the ELCS-K data will be analyzed.

Other information related to demographic characteristics (e.g., children's race/ethnicity, household poverty status, parents' highest level of education, family type, and primary home language) was collected via interviews with parents during the first wave of the investigation. Most of the parent or guardian interviews were conducted by telephone or in person. The respondent to the parent interview was the individual who identified as knowing the most about the child's care, education, and health. In some instances, the parent interview was fully translated into Spanish before data collection began and could be administered by bilingual interviewers if parent respondents preferred the interview in Spanish. Because it was cost prohibitive to do so, the parent interview was not translated to other languages. In those cases, parents who spoke other languages were interviewed using an interpreter.

Measures and Procedures

Dependent Variables

Because the focus of this investigation is mathematics achievement, this assessment will be used as the outcome or dependent variable. Mathematics achievement is a unidimensional assessment that included topic areas such as number properties and operations, measurement, geometry, data analysis and probability, and algebra. As the participants got older, questions from more advanced topics increased in number and more basic questions decreased. IRT based scores were calculated in the form of theta scores that range from -8 through 8 and measure how children perform in different mathematics topics. Theta scores were equated over age/grade to provide a developmental scale that can be used to measure gains over time.

Independent Variables

As a predictor of mathematics achievement, EF is composed of several assessments that measure the different domains. Participants were tested for the cognitive flexibility component of working memory. Demographic characteristics were also used to predict mathematics achievement. The literature and critical current work have informed the use of these variables in understanding mathematics achievement. Note that inhibitory control is not considered further since it was not measured throughout the entire length of the investigation.

Cognitive Flexibility. An individual's ability to see a task or aspect of a problem from differing perspectives is a useful skill and refers to what researchers call cognitive flexibility (Diamond, 2013). Cognitive flexibility is more commonly known as an individual's ability to 'think outside the box' and is a skill that is widely assessed by researchers who study children's academic achievement (Zelazo, 2015). This skill is frequently investigated by researchers in a wide array of contexts, such as task switching and set-shifting tasks (Diamond, 2013). The Dimensional Change Card Sort (DCCS) will be used throughout this investigation to assess children's cognitive flexibility. The task will allow researchers to determine whether cognitive flexibility measures can predict a child's mathematical achievement.

Cognitive flexibility was measured throughout the entire investigation (e.g., Kindergarten through 5th grade). The assessment used to measure this domain was the dimensional change card sort (DCCS) from the NIH Toolbox. This task involves sorting pictures such as a sailboat or rabbit. The pictures are sorted into one or two categories consistent with the rules provided. This investigation will use scores for the total or entire assessment. During the first two years (e.g., Kindergarten through 1st grade) of the investigation, participants were given the physical version of the DCCS. The combined accuracy for the physical version reflects accuracy and ranges between 0-18, where a higher score indicates greater cognitive flexibility. A computerized version of the DCCS was created for second graders by the National Institute of Health Toolbox for the Assessment of Neurological and Behavioral Function (NIH Toolbox). The computerized version of the DCCS makes the test age appropriate for older children (Tourangeau, Nord, Le & Wallner-Allen, 2017). However, it is not seen as an age-appropriate tool for participants in the first grade and younger. Only children aged eight and older were given the computerized version of the DCCS that was used for later data collection waves (e.g., 2nd through 5th grade). For this version of the assessment, overall score ranges were between 0-10, with a higher score indicating greater accuracy and reaction time. It is acknowledged that the DCCS was scaled differently for different grade years (e.g., Kindergarten through 1st grade and 2nd through 5th grade) with the same population across the duration of the investigation. Because of this change in scaling, the analysis using DCCS scores will be limited to the period beginning in Fall of the second grade.

As an instrument of cognitive flexibility, the DCCS is appropriate for subjects of 3-85 years of age. To make the assessment age appropriate, the DCCS is administered at different starting points with 40 different trials. The assessment starts with the participant sorting pictures based on a given dimension (e.g., color). This starting point is termed the pre-switch trial and the first part of the test consists of five such trials. Subsequently, there are five trials (post-switch trials) where the participant sorts items based on different dimensions (e.g. shape). The complexity of the assessment starts when the participant is asked to sort the pictures using different dimensions (e.g., color or shape) with a total of 30 mix trials. To account for age, participants under age eight start the assessment with pre-switch trials and participants over age 8 start with the mix block trials. Equal points are given for correctly applying the rule at the respective trial (e.g., pre-switch or post-switch trials). When obtaining the final score from the physical and computerized version of the DCCS, it is important to note that the computerized version allows for more accurate time information, as it captures trial data in milliseconds. Thus, accuracy, and not time, will be the best indicator of cognitive flexibility for the DCCS data (Tourangeau, Nord, Le & Wallner-Allen, 2017).

Administering the DCCS requires the administrator to determine that the participant can demonstrate an understanding of the task, which consists of 8-24 trials. The test administrator instructs participant on the identification rules for the pictures (e.g., color or shape). Four practice trials are administered with the test administrator's instructions. During these beginning trials, the stimulus is sorted by shape. Successfully sorting 3 out of 4 trials will lead to practice sorting stimulus according to color. If the participant does not apply the rule correctly, then another practice round will be administered. Only if the participant successfully completes the trials will they be administered the DCCS (Tourangeau, Nord, Le & Wallner-Allen, 2017). The actual scored DCCS is composed of 30 separate trials. A scoring algorithm obtained from the NIH Toolbox was used to obtain the scores. The final score for the DCCS ranges between zero and ten. The participants' score accuracy (0 to 5 units) and reaction time (0 to 5 units) factored into the final scores (Tourangeau, Nord, Le & Wallner-Allen, 2017, p. 3-17). Accuracy that is at 80 percent or less will lead to an overall score based on the participant's accuracy. Accuracy that is greater than 80 percent results in an overall score that combines both reaction time and rule application accuracy (Tourangeau, Nord, Le & Wallner-Allen, 2017). The overall accuracy score can range between 0 to 5. The forty outcomes of 40 trials was reduced to a maximum score of five. This means that the participant earns a total of 0.125 points (5 points/40 trials) for each correct trial. In the event that only 30 mixed block trials are administered, five pre-trial points and five post-trial points (a total of 10 points), the overall accuracy for the DCCS is computed as follows:

*DCCS accuracy score = .125 * # of correct responses*

Working Memory. Working memory, a component of EF, is a cognitive function that is composed of the ability to retain and manipulate information. This domain is an ideal measure for researchers to assess executive function and academic achievement. Working memory has been previously used to predict academic achievement. Working memory was assessed throughout the entire investigation (e.g., Kindergarten through 5th grade), via the Woodcock-Johnson III numbers reversed subtest. The test yields four scores (e.g., age standard score, age percentile score, grade standard score, and grade percentile score). The scores indicate performance relative to peers' grade-normed scores and age-normed transformation data. For this test, the participant is shown a series of numbers and asked to repeat the numbers, in reverse order, with increasing difficulty or length. For this investigation, the pre-calculated age standard score will be used in the analysis. The standard score reflects the participant's performance in comparison to age and grade level average, with a mean score of 100 and a standard deviation of 15 (Tourangeau, Nord, Le & Wallner-Allen, 2017, p. 3-28). The numbers reversed subtest scores were also computed as z-scores using the participant's *W* score, with a z-score having a mean of zero and a standard deviation of one (Tourangeau, Nord, Le & Wallner-Allen, 2017, p. 3-28). Z-scores that were standardized across grade level (e.g., Kindergarten through 5th grade) were used in this investigation.

Because the numbers reversed subtest yields different outcome scores, the preferred outcome score can vary depending on the investigation. Grade or age percentile rank or standardized scores may be better suited for analyzing data for one time-point. A *W* score is better suited for longitudinal analysis. The *W* score is a standardized score that is produced from the Rasch ability scale (Tourangeau, Nord, Le & Wallner-Allen, 2017, p. 3-25). It provides a common scale that represents both difficulty and ability of the task. The *W* scale is most useful for measuring growth. It has a mean of 500 and standard deviation of 100. The mean average of the WJ III is set for a child at 10 years of age. Participants younger than 10 years are expected to have a *W* score lower than 500, while older children should have a *W* score greater than 500. The *W* score is expected to increase as the child develops with age. A child's *W* score increasing from 430-440 indicates growth. The *W* score is an equal-interval scale score that is suited for both regression and correlation analyses. The higher the *W* score, the better they performed at reiterating more difficult number sequences (Tourangeau, Nord, Le & Wallner-Allen, 2017, p. 3-26). However, the *W* score does not reflect the child's pattern of responses; it only reflects the total number of correct answers (p. 3-26). Thus, the *W* score is not an indication of the number of sequences the child correctly answers (Tourangeau, Nord, Le & Wallner-Allen, 2017).

The Numbers Reversed subtest was created to be administered during early childhood education (e.g., K-5th grade) (Tourangeau, Nord, Le & Wallner-Allen, 2017) and is meant to tap into the participant's working memory. The subtest consists of the backward digit span task. The backward digit span task consists of the participant having to reiterate in reverse order a sequence of the number shown (Tourangeau, Nord, Le & Wallner-Allen, 2017, p. 3-24). If a child was presented with a series of numbers (e.g., "3…5"), the correct answer would be stated in reverse order (e.g., "5…3"). Participants were administered the Numbers Reversed task for two-number sequences five times. The Numbers Reversed task ends when three consecutive number mistakes are made. Individuals who answered correctly were given up to five, three-number, sequences. The sequences then become increasing longer, up to a maximum of eight numbers, until the participant makes three consecutive mistakes (Tourangeau, Nord, Le & Wallner-Allen, 2017).

The average score for the Numbers Reversed subtest is 403 for children who have not yet developed their working memory skill. As the child's working memory develops, the child's WJIII Numbers Reversed subtest measures can be compared to the child's baseline *W* score. The participant's fall and/or spring kindergarten *W* score can then be compared to other participants' scores to determine growth in future assessments (Tourangeau, Nord, Le & Wallner-Allen, 2017, p. 3-27). Researchers are cautioned when interpreting the raw score of zero (which is a 403 in *W* score) because it is not an accurate measure of a child's working memory ability. For example, in one study that administered the Numbers Reversed subtest to kindergarteners in the Fall semester, roughly forty percent of the participants had not developed their working memory skills and were therefore given a score of 403 (Tourangeau, Nord, Le & Wallner-Allen, 2017, p. 3-27). As the participant's skills developed, the following Spring semester, only twenty percent of the students were not able to score above the lowest score scalable. The following academic year, fall semester, when participants were in the first grade, that number had dropped to less than thirteen percent for children who scored the lowest score scalable. Only six percent of children scored the lowest during the Spring semester of their first-grade year. It is not clear, on the other hand, how close some of these children were to scoring higher on the Number Reversed subtest (Tourangeau, Nord, Le & Wallner-Allen, 2017).

One factor contributing to children scoring 403 or less (0 in raw score) is the fact that the practice items were improperly administered by the assessors, resulting in some of the children not fully understanding specific aspects of the Numbers Reversed subtest. During field observations in one study, it was noted that there were inconsistencies with how the practice items were administered when the children did not answer the first practice item correctly. However, it is difficult to determine to what extent the practice items being improperly administered to the children affected the results of the study. Consequently, proper administration of practice items may impact the performance for some of the participant's score. It is therefore important for researchers to decide how to handle scores of 403 or less during analyses. This decision should be based on the analytical goals of both the researcher and the study.

Demographic Characteristics

Gender. Information about the child's gender was collected from the respective schools at the time of sampling. The information was stored in the study's administrative database and collected again from the parents in the fall kindergarten parent interview. This information was further confirmed by parents in the spring kindergarten parent interview and then asked again in the later round of interviews if the data were either missing or had never been confirmed by the parent (Tourangeau et al. 2017).

Race. Information about the participants' race was collected from the parents as part of the parent interview. This parent information was cross-checked with the school to ensure consistency. The parents provided the participants' race during the kindergarten academic year (2010–11). In the spring of 2012, the spring of 2013, the spring of 2014, and the spring of 2015, parents were asked to provide information on the child's ethnicity and race in case the information was missing or had not been confirmed by a parent/guardian in a previous data collection wave (Tourangeau et al. 2017).

During the parent interview, respondents were asked to indicate whether their child belonged to one or more of the following races: White, Black or African American, American

Indian or Alaska Native, Asian, and Native Hawaiian or Other Pacific Islander. In addition, each parent was asked to identify whether the child was Hispanic or Latino. Hispanic and race were used to create eight mutually exclusive categories: White, not Hispanic; Black or African American, not Hispanic; Hispanic, race specified; Hispanic, no race specified; Asian, not Hispanic; Native Hawaiian or Other Pacific Islander, not Hispanic; American Indian or Alaska Native, not Hispanic; and Two or more races, not Hispanic. For this study, a race variable has been created with the categories of African American, Asian American, White, Hispanic, and Other.

Socioeconomic Status. Socio-economic status is defined based on a government assigned classification. One critical classification is poverty status. Poverty status was created based on household income and the total number of household members. To get this information, the parents of the participants were asked to report their household income based on a standard list of categories. A classification of poverty was given based on exact income and household size. This was done using the weighted 2010 census poverty thresholds. A classification below the federal poverty level is a two person household that earns an income below $14,220. If the same household earns more than $14,220 but less than $28,440, then their classification is at the poverty level. A household above an income of $28,440 was given a classification above the poverty level (Mulligan, McCarroll, Flanagan, & Potter, 2014).

Statistical Analysis

A series of descriptive statistical analyses will be conducted on variables of interest in order to present an initial illustration of the ECLES- K: 2011 dataset used for this study. First, frequency tables will be constructed for all categorical measures included in this investigation, with the sample sizes and percentages of response reported for each response category (Holcomb, 2016). Following this, means and standard deviations will be calculated and reported for all continuous measures of interest included within this study. The descriptive statistics will then be followed by fitting models to address the study's research questions. Summary statistics for time-varying variables will be presented for each measurement period in the data.

To review, the research questions of interest for this investigation consist of the following:

Research Question 1: *What is the average math achievement growth trajectory from kindergarten through fourth grade, accounting for the nesting of students within schools, and summer loss?*

Research Question 2: *To what degree do gender, racial, or socioeconomic status differences account for variation in the intercepts, slopes, and summer drops in the mathematics trajectories at the student and school levels?*

Research Question 3: *Is executive functioning associated with change in math achievement over this period?*

Research Question 4: *Which student sociodemographic characteristics are most strongly associated change in math achievement, and to what degree does controlling for student sociodemographic characteristics alter the association between executive functioning and change in math achievement?*

Research Question 5. *To what degree does math achievement growth vary among schools, controlling for demographics?*

In all cases, linear mixed effects models (LME) will be used to account for the nested structure of the data. The first source of clustering is due to the repeated measures (fall of kindergarten through spring of fourth grade) nested within students. The second level of clustering is due to students nested within schools. LME modeling allows researchers to account for the nested structure of the data, including the nesting of time points within cases (Raudenbush & Bryk, 2002). LME is preferred to traditional regression modeling because of the fact that observations are not treated as independent and standard error estimates are less biased, thus producing more precise inferences than other approaches.

The LME models to be tested will include random effects for the subject and school. In addition, the effect of time will be treated as randomly varying between students. This yields a more complicated error structure that no longer assumes zero correlations between cases. When the effect of time is allowed to be individual-specific, the within-individual error covariances are not only non-zero but also a function of time (Fitzmaurice, Laird & Ware, 2011). In addition, LME models are more robust to missing data than AN(C)OVA models because a missing time point does not require dropping all the observed time points for that individual (Raudenbush & Bryk, 2002).

Because of the large sample size and number of random effects in the models, the LME models will be estimated using the hpmixed procedure in SAS v. 9.3. The first measurement for each subject will be coded as one to indicate baseline, which, on a log scale, becomes zero. All subsequent measures will be coded in increments of 1/12. That is, time will be coded as follows:

- GradeK-September = 0/12
- GradeK-October = 1/12
- GradeK-November = 2/12
- GradeK-December = 3/12
- GradeK-March = 6/12
- GradeK-April = 7/12
- GradeK-May = 8/12
- Grade1-September = 12/12
- Grade1-October = 13/12
- …
- Grade4-May = 56/12

An additional variable, *drop*, will be incorporated into the models that coded as zero for Grade K, 1 for Grade 1, 2 for Grade 2, etc. The purpose of this variable is to introduce a loss upon grade transition (summer), with the assumption the amount of loss is equal at each transition.

The research questions will be addressed using the following models:

Research Question 1: *What is the average math achievement growth trajectory from kindergarten through fourth grade, accounting for the nesting of students within schools, and summer loss?*

The first research question only considers the typical trajectory of math achievement after accounting for student and school random effects. This model will be used to establish baseline, unadjusted trends against which subsequent models can be compared.

The model that will be fit is the following:

$$MA_{tis} = b_0 + b_1 ln(time)_{tis} + b_2 drop_{tis} + +u_{0s} + u_{0is} + u_{1s} ln(time)_{tis} + u_{1is} ln(time)_{tis} + u_{2s} drop_{tis} + u_{2is} drop_{tis} + r_{tis}$$

Here, MA_{tis} is the value of math achievement at time *t* for the *i*th student in school *s*. b_1 is the value of the average linear trajectory (after a natural log transform), and b_2 represents the dip in the fall due to summer loss. u_{0s} is the school random effect, u_{1is} is the student (nested in school) random effect, u_{1s} is the school-level random effect for time, u_{1is} is the student-level random effect for time, u_{2s} is the school-level random effect of summer drop, and u_{2is} is the student-level effect for random drop. The time and drop random effects account for differences in trends and summer drop between schools and students. The final term is remaining error. The random effects matrices at the school and student levels will be treated as unstructured, allowing covariances between random effects to be estimated, and independent between levels.

As will be discussed in greater detail in the next chapter, the choice of using a natural log transformation for time is based on two considerations. First, although a quadratic functional form is a common way to capture nonlinear trends, it introduces the complications in both interpretation and estimation due to the requirement of an additional random effect at both the student and school levels. Second, based on graphical evaluations of trends, the log transform appears to accurately capture nonlinearities. Although in theory there may be student and school-level differences in the (log) trajectories and summer drop effects, estimation problems can occur if estimates of the variance components approach zero. This is often manifest in random effects correlations near unity. Hence, the results will report models that constrain some of the variance components to zero along with change in model fit statistics so that the reader can see the implications for overall inferences of fixed effects when changing assumptions about the random effects.

Research Question 2: *To what degree do gender, racial, or socioeconomic status differences account for variation in the intercepts, slopes, and summer drops in the mathematics trajectories at the student and school levels?*

This second research question focuses on whether there are gender, racial, and socioeconomic status differences in mathematics achievement performance. In order to test these associations, an MLM will be conducted which will include these demographic measures as independent variables and mathematics achievement performance as the dependent variable, with the data being clustered within subject and school. The results of this analysis will determine whether there are any significant differences in mathematics performance as a result of these demographic measures. The model will be the following:

$$\begin{aligned} MA_{tis} = {} & b_0 + b_1 ln(time)_{tis} + b_2 drop_{tis} + b_3 gender_i + b_4 black_i + b_5 hispanic_i + b_6 asian_i \\ & + b_7 other\ race_i + b_8 ses_i + b_9 gender_i ln(time)_{tis} + b_{10} black_i ln(time)_{tis} \\ & + b_{11} hispanic_i ln(time)_{tis} + b_{12} asian_i ln(time)_{tis} \\ & + b_{13} other\ race_i ln(time)_{tis} + b_{14} ses_i ln(time)_{tis} + b_{15} gender_i drop_{tis} \\ & + b_{16} black_i drop_{tis} + b_{17} hispanic_i drop_{tis} + b_{18} asian_i drop_{tis} \\ & + b_{19} other\ race_i drop_{tis} + b_{20} ses_i drop_{tis} + u_{0s} + u_{0is} + u_{1s} ln(time)_{tis} \\ & + u_{1is} ln(time)_{tis} + u_{2s} drop_{tis} + u_{2is} drop_{tis} + r_{tis} \end{aligned}$$

Coefficients b_3 through b_8 represent the main effects of each demographics variable, or the effect when time = 0 (fall of kindergarten year). Coefficients b_9 through b_{14} are the interactions with time, which estimate how much the trajectories vary as a function of each demographics variable. Coefficients b_{15} through b_{20} are the interactions with summer drop, which estimate how much the effect of summer drop varies as a function of each demographics variable. The random effects u_{0s} through u_{2is} summarize how much the intercept, trajectory, and summer drop vary at the student and school levels net of any variability captured by the demographics.

When an interaction is significant, the fixed effect will be examined graphically by plotting the trajectory for each value of the demographic variable. This will facilitate interpretation given that, with a logged time variable, the rate of change will depend on the time point.

Research Question 3*: Research Question 3: Is executive functioning associated with change in math achievement over this period?*

The analysis of the third research question research question will progress sequentially by first modeling EF and then correlating the predicted random effects of the EF model with the predicted random effects of the math achievement model. The following model will be fit to working memory:

$$WM_{tis} = b_0 + b_1 ln(time)_{tis} + b_2 drop_{tis} + u_{0s} + u_{0is} + u_{1s} ln(time)_{tis} + u_{1is} ln(time)_{tis} + u_{2s} drop_{tis} + u_{2is} drop_{tis} + r_{tis}$$

The model terms are interpreted in the same manner as those in the model of math achievement described for research question one. However, whereas the first two research questions focused on fixed effects and variance components (along with their covariances), the third research question is interested in knowing if the student-specific and school-specific variability in EF is associated with the student-specific and school-specific variability in math achievement. To determine if this is the case, the best linear unbiased predictions (BLUPs) of the student and school specific random effects will be estimated and saved for both EF and math achievement. They will then be combined into a single data file, and a correlation matrix of predicted values will be estimated.

Specifically, correlation between the u_{0is} term from the EF equation and the u_{1is} term from the math equation will indicate if early EF is associated with larger math gains among students. This represents the association most directly related to the research question. In addition, the other random effects correlations at both the student level and school level will also be presented for completeness. The correlation between the predictions from the two u_{0is} terms (one from each model) will indicate the size of the association between initial levels of math and EF at the student level (do kindergartners with high EF tend to have higher math achievement?). The correlation between the predictions of the two u_{1is} terms will indicate the size of the association between the EF and math trajectories at the student level (do increases in EF tend to go along with increases in math achievement?). The correlation between the predictions of the two u_{2is} terms will indicate the size of the association between the EF and math summer drop effects at the student level (do increases in drop in EF tend to go along with increases in drop in math achievement?). The correlation between the u_{0is} term from the math equation and the u_{1is} term from the EF equation will indicate if early math achievement is associated with larger gains in EF.

The same correlation matrix will be assessed at the school level. Specifically, the correlation between the predictions from the two u_{0s} terms (one from each model) will indicate the size of the association between initial levels of math and EF at the school level (do schools with high EF in kindergarten tend to have higher math achievement in kindergarten?). The correlation between the predictions of the two u_{1s} terms will indicate the size of the association between the EF and math trajectories at the school level (do schools demonstrating increases in EF tend to also have students demonstrating increases in math achievement?). The correlation between the predictions of the two u_{2s} terms will indicate the size of the association between the EF and math summer drop effects at the school level (do schools with larger summer drop in EF also tend to have larger summer drop effects in math achievement?). The correlation between the u_{0s} term from the EF equation and the u_{1s} term from the math equation will indicate if early EF is associated with larger math gains between schools. The correlation between the u_{0s} term from the math equation and the u_{1i} term from the EF equation will indicate if schools having higher early math achievement tend to show larger gains in EF.

This process will be repeated using DCCS scores. However, because the scaling of DCCS changed in the second grade, the model will only be fit using data starting from grade two. In addition, since most of the subsequent data collection occurred in the spring, there will be no summer drop variable included in the model. The specification will be the following:

$$DCCS_{tis} = b_0 + b_1 ln(time)_{tis} + u_{0s} + u_{0is} + u_{1s} ln(time)_{tis} + u_{1is} ln(time)_{tis} + u_{0s} + u_{0is} + u_{1s} ln(time)_{tis} + u_{1is} ln(time)_{tis} + r_{tis}$$

For these models, time will be rescaled so that second grade equals 1 (zero on the log scale). As with working memory, the random effects will be predicted and then correlated with the random effects from a math achievement model fit to the same period.

Research Question 4: *Which student sociodemographic characteristics are most strongly associated change in math achievement and to what degree does controlling for student sociodemographic characteristics alter the association between executive functioning and change in math achievement?*

The final research question is similar to the last except that the models now incorporate demographics. For working memory, the model of interest is the following:

$$WM_{tis} = b_0 + b_1 ln(time)_{tis} + b_2 drop_{tis} + b_3 gender_i + b_4 black_i + b_5 hispanic_i + b_6 asian_i + b_7 other\ race_i + b_8 ses_i + b_9 gender_i ln(time)_{tis} + b_{10} black_i ln(time)_{tis} + b_{11} hispanic_i ln(time)_{tis} + b_{12} asian_i ln(time)_{tis} + b_{13} other\ race_i ln(time)_{tis} + b_{14} ses_i ln(time)_{tis} + b_{15} gender_i drop_{tis} + b_{16} black_i drop_{tis} + b_{17} hispanic_i drop_{tis} + b_{18} asian_i drop_{tis} + b_{19} other\ race_i drop_{tis} + b_{20} ses_i drop_{tis} + u_{0s} + u_{0is} + u_{1s} ln(time)_{tis} + u_{1is} ln(time)_{tis} + u_{2s} drop_{tis} + u_{2is} drop_{tis} + r_{tis}$$

The DCCS model, fit only to the data beginning in second grade, will be the following:

$$DCCS_{tis} = b_0 + b_1 ln(time)_{tis} + b_2 drop_{tis} + b_3 gender_i + b_4 black_i + b_5 hispanic_i + b_6 asian_i + b_7 other\ race_i + b_8 ses_i + b_9 gender_i ln(time)_{tis} + b_{10} black_i ln(time)_{tis} + b_{11} hispanic_i ln(time)_{tis} + b_{12} asian_i ln(time)_{tis} + b_{13} other\ race_i ln(time)_{tis} + b_{14} ses_i ln(time)_{tis} + u_{0s} + u_{0is} + u_{1s} ln(time)_{tis} + u_{1is} ln(time)_{tis} + r_{tis}$$

The focus of interpretation for both outcomes will again be on the correlations between the random effects predictions with the research question addressed most directly by the correlation between the EF scores in kindergarten and the math achievement trajectory. The primary difference from the prior research question is that the random effects are now estimated net of any variance explained by the inclusion of the demographics variables.

In a supplementary analysis, a model of math achievement will be fit that includes both demographics and the EF measures (separate models for working memory and cognitive flexibility). The purpose will be to determine if EF adds significantly to the explanatory power of the model. Whereas the other models corresponding to research question four are interested in the association between intercepts and trajectories, these final models will assess how much EF directly affects math achievement. The model including working memory will be the following:

$$MA_{tis} = b_0 + b_1 ln(time)_{tis} + b_2 drop_{tis} + b_3 gender_i + b_4 black_i + b_5 hispanic_i + b_6 asian_i + b_7 other\ race_i + b_8 ses_i + b_9 gender_i ln(time)_{tis} + b_{10} black_i ln(time)_{tis} + b_{11} hispanic_i ln(time)_{tis} + b_{12} asian_i ln(time)_{tis} + b_{13} other\ race_i ln(time)_{tis} + b_{14} ses_i ln(time)_{tis} + b_{15} gender_i drop_{tis} + b_{16} black_i drop_{tis} + b_{17} hispanic_i drop_{tis} + b_{18} asian_i drop_{tis} + b_{19} other\ race_i drop_{tis} + b_{20} ses_i drop_{tis} + b_{21} working\ memory_{tis} + u_{0s} + u_{0is} + u_{1s} ln(time)_{tis} + u_{1is} ln(time)_{tis} + u_{2s} drop_{tis} + u_{2is} drop_{tis} + r_{tis}$$

The focus of interpretation will be on b_{21}, the coefficient corresponding to the effect of working memory. In addition, the model fit will be compared to what was obtained in the model that excludes the working memory fixed effect.

For DCCS, the model will be the following:

$$MA_{tis} = b_0 + b_1 ln(time)_{tis} + b_2 drop_{tis} + b_3 gender_i + b_4 black_i + b_5 hispanic_i + b_6 asian_i + b_7 other\ race_i + b_8 ses_i + b_9 gender_i ln(time)_{tis} + b_{10} black_i ln(time)_{tis} + b_{11} hispanic_i ln(time)_{tis} + b_{12} asian_i ln(time)_{tis} + b_{13} other\ race_i ln(time)_{tis} + b_{14} ses_i ln(time)_{tis} + b_{15} DCCS_{tis} + u_{0s} + u_{0is} + u_{1s} ln(time)_{tis} + u_{1is} ln(time)_{tis} + r_{tis}$$

Here the focus of interpretation will be the estimate of b_{15}, the fixed effect for cognitive flexibility. The model fit will be compared against a model of math achievement estimated without the DCCS variable using the same sample (beginning in second grade).

Research Question 5. *To what degree does math achievement growth vary among schools, controlling for demographics?*

For each of the models corresponding to research questions 3 and 4, a table will be presented showing the amount of variability explained at both the student and school levels by inclusion of demographics. The final research question returns to these tables and focuses specifically on how much between-school variability remains. In addition, boxplots of the predicted random effects will be presented for the models with and without demographics in order to visualize the reduction in variability. This information helps to understand how much school quality matters for math achievement outcomes after taking into account student characteristics that cannot be targeted by policy.

Additional Model Considerations

Assumptions and Diagnostics. LME makes the assumption that, conditional on the fixed and random effects, the individual-level variance is normally distributed with a mean of zero and constant variance. It also assumes that the variance components at each level are distributed multivariate normal and uncorrelated between levels. Another assumption is that the relationships between independent variables and the outcome are linear in the coefficients and that outliers do not have a disproportionate effect on the results.

The first assumption will be tested using a q-q plot of the level-1 residuals, where the residuals should follow a 45^o line. A plot of standardized residuals by predicted values will be used to assess for non-constant variance, with evidence of heteroskedasticity occurring if the spread of the residuals gets larger or smaller across the range of predicted values. Boxplots of residuals by school will also be examined to determine if there are systematic differences in variance by the highest level of grouping.

The appropriateness of the functional form (linearity) will be assessed by looking at plots of each independent variable (x-axis) and the outcome (y-axis) and fitting a *loess* line to the data. A *loess* line, unlike a linear regression line, is allowed to curve and follow the density of the data. If the line shows substantial curve, the variable will be transformed appropriately, either using a log or polynomial transformation.

The sample size will be large, so it is unlikely that outliers will have a large influence on the results. Nonetheless, standardized individual-level residuals will be calculated, and cases with a value greater than +/- 4 will be investigated to determine if they should be included or excluded.

CHAPTER 4

RESULTS

Tables 1 and 2 present descriptive statistics for the main study variables. Separate descriptive statistics are presented for the unweighted sample, the sample after applying the survey weights, and finally the sample with survey weights after imputing missing values. Missing data imputation was carried out in SPSS v. 25. Imputation was done in “wide format”, i.e., with a single record per child and repeated measures over time represented as distinct variables (e.g., ma.1 to ma.8 to represent math at the eight waves of data collection). Thus, the imputation accounted for correlations among the repeated measures over time within person (dependence within child). The imputation model included repeated measures on math, working memory and dccs (total scores prior to grade 2 and scores using RT from grade 2 onwards), as well as race, gender, and SES. In addition, to also account for dependence within schools, school means were calculated for math, working memory, DCCS total scores (< grade 2), and DCCS overall scores w/ RT information (grade 2+) and these were included in the imputation model. This was done because otherwise failing to include information on between-school differences would lead to artificial deflation of the school-level variance component. The student-level weights were re-scaled to sum to the number of children within the school, as recommended in the Stata manual.

The descriptive statistics in Tables 1 and 2 use the original, raw survey weights. Table 1 contains the variables measured on an interval scale and that, apart from SES, vary over time. Table 2 shows that the weights were more consequential for the demographics, but imputation does not have much effect on the distribution of race or gender. To maximize power, the model results reported in this chapter are based on the weighted and imputed values.

Table 1

Descriptive Statistics for Continuous Variables

Variable	Grade	Semester	N Observed	N with valid weights	N with valid weights and impute	Unweighted Mean	Weighted Mean	Weighted Mean with impute	SD
Mathematics	K	F	15,222	13,883	14090	-0.49	-0.50	-0.51	0.93
(theta scores)	K	S	17,128	13,859	14090	0.45	0.43	0.43	0.77
	1	F	5,215	4,063	14090	0.89	0.89	0.92	0.83
	1	S	15,090	12,103	14090	1.65	1.65	1.63	0.85
	2	F	4,720	3,683	14090	1.86	1.86	1.88	0.85
	2	S	13,811	11,117	14090	2.45	2.45	2.42	0.82
	3	S	12,845	10,342	14090	3.05	3.05	3.01	0.77
	4	S	12,060	9,711	14090	3.42	3.42	3.38	0.78
Working Memory	K	F	15,227	13,888	14090	433.13	433.21	433.01	30.26
(W scores)	K	S	17,132	13,867	14090	449.68	450.01	449.90	30.52
	1	F	5,215	4,064	14090	456.97	457.50	458.33	28.73
	1	S	15,094	12,106	14090	469.33	469.50	469.09	25.82
	2	F	4,718	3,682	14090	473.25	473.45	473.54	24.34
	2	S	13,813	11,118	14090	480.70	480.63	480.01	23.26
	3	S	12,856	10,346	14090	489.82	489.90	489.32	22.21
	4	S	12,065	9,712	14090	497.26	497.28	496.73	21.69
DCCS	2	F	4,699	3,668	14090	6.30	6.34	6.30	1.47
(overall)	2	S	13,755	11,069	14090	6.69	6.71	6.67	1.35
	3	S	12,723	10,235	14090	7.19	7.22	7.19	1.11
	4	S	12,001	9,656	14090	7.62	7.63	7.62	0.98
SES	K-1	NA	16,543	13,251	14090	-0.07	-0.06	-0.07	0.82

Note: Excludes invalid school ID, assessments made in June-August

Table 2

Descriptive Statistics - Categorical Variables

	Unweighted		Weighted (valid weights)		Weighted & Imputed	
	Freq	%	Freq	%	Freq	%
Race						
Asian	1,646	9.27	169,722	4.71	169,982	4.71
Black	2,308	13	471,705	13.10	474,611	13.15
Hispanic	4,525	25.49	865,194	24.03	866,876	24.02
Other	964	5.43	189,845	5.27	190,668	5.28
White	8,308	46.8	1,903,393	52.87	1,906,218	52.83
Gender						
Male	9,095	51.22	1,850,587	51.39	1,853,919	51.38
Female	8,661	48.78	1,750,693	48.61	1,754,436	48.62

Note: Excludes invalid school ID, assessments made in June-August set to missing.

Modeling the Primary Outcomes

The first step in the analysis is to determine the appropriate functional forms for the student trajectories. Figure 1 presents three panels summarizing the math achievement trajectories for a random sample of 50 students (a larger number would make the graph difficult to read). The scores have been standardized to have zero mean and unit variance due to the fact that the three primary outcomes (math achievement, working memory, and cognitive flexibility) are measured on very different scales. Having a common metric will facilitate comparing random effects between outcomes later in the chapter.

The first panel of Figure 1 shows the observed scores at the different time points connected by a separate line for each student. There appears to be nonlinearity in the trends, with somewhat steeper slopes at the beginning of the series that level off in later years. There are two common ways to model this type of nonlinearity, by specifying time as a second-order polynomial or by using a base-*e* log transform of the time variable. The second panel shows the lines modeled using the quadratic specification, and the final panel shows the natural log version. Visually, both approaches appear to be capturing the nonlinear trend.

There are practical reasons to prefer the log version to the polynomial. The polynomial requires two terms to be fit, the linear parameter (time) and the acceleration parameter (time squared). This means that it will be necessary to include both terms in the student-level and school-level random effects, which increases the number of variance components and covariances to be estimated at each level. Attempts to fit the polynomial specification in Stata and R (lme4) using all of the data resulted in non-convergence or singular solutions, only SAS's proc hpmixed was able to produce results (after nearly an hour of run time for the model *without* demographics). In addition, interpretation of the random effects becomes more difficult, especially when comparing random effects between outcomes. This is because the association between the math achievement and executive function trajectories need to be summarized with two separate correlations, and neither time nor time-squared by themselves tell the complete story. For these reasons, and because the natural log transformation still captures the observed nonlinearity, the polynomial specification will not be used.

Figure 2 presents similar graphs to Figure 1 but for working memory. The individual trajectories are a little messier compared to math achievement. However, the polynomial and natural log fits show that the typical trajectory is again nonlinear. The same caveats about estimation and interpretation complications exist for the polynomial specification. Again, as both the log and polynomial models capture diminishing returns over time, and because the natural log is easier to fit and interpret, the quadratic approach will not be used.

Figure 3 shows the trends for cognitive flexibility, with measures beginning in the second grade and no fall measures in the third or fourth grades. The polynomial and log fits both produce relatively linear trends. However, to be consistent with the prior outcomes, the natural log version will be used. This puts the time variable on the same scale for math achievement and cognitive flexibility when comparing their respective random effects.

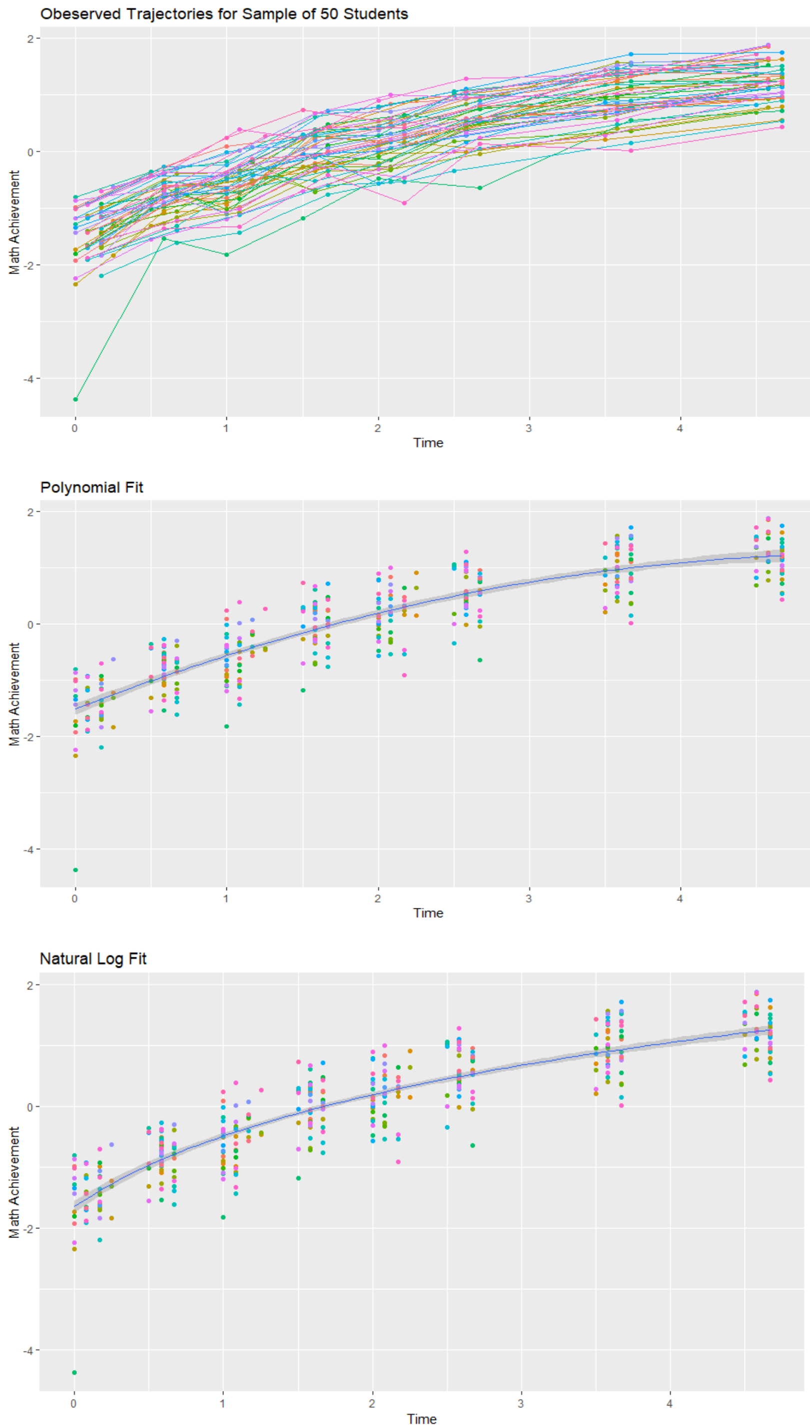

Figure 1: Plots of math achievement trajectories for random subsample of 50 students. First panel is the observed trajectories. Second panel is the polynomial fit to the subsample. Third panel is the natural log fit to the subsample.

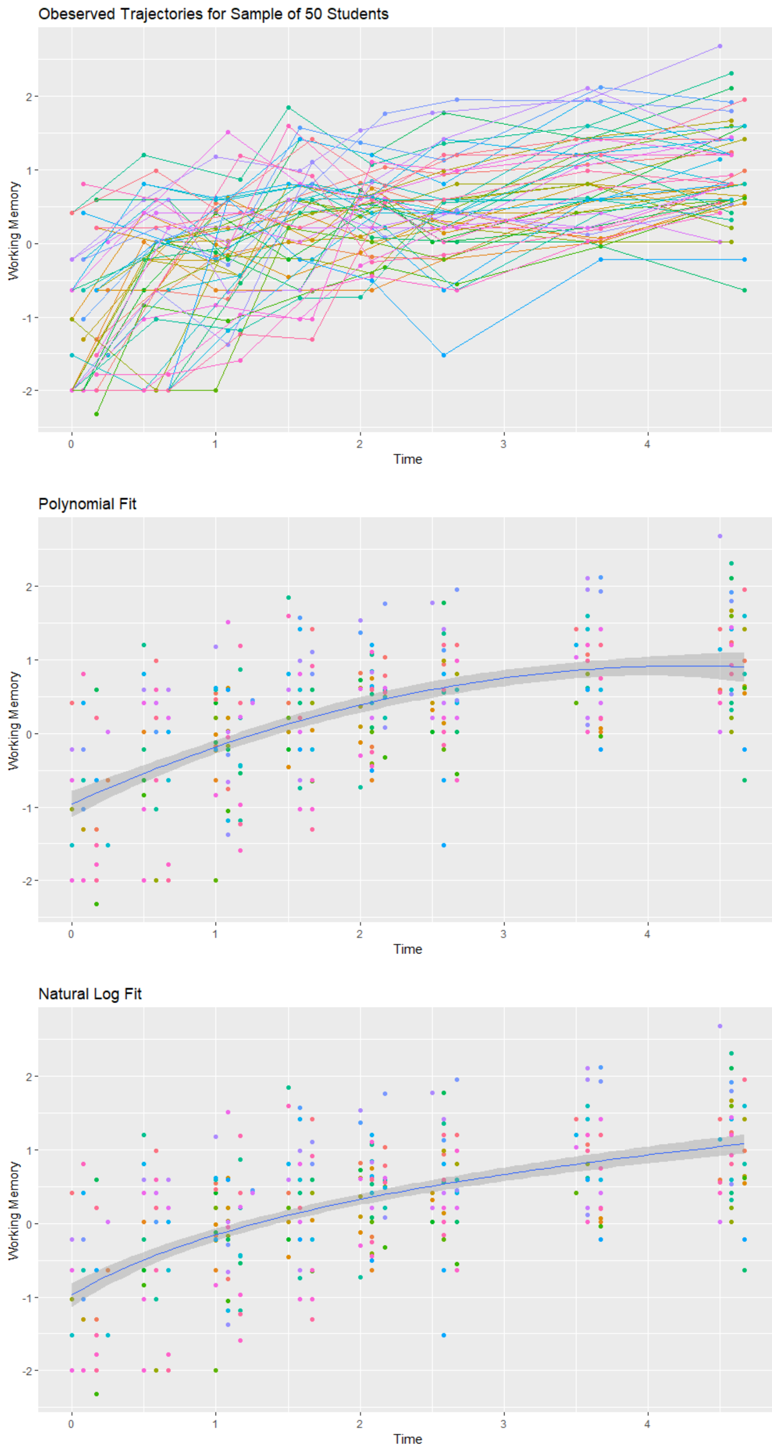

Figure 2: Plots of working memory trajectories for random subsample of 50 students. First panel is the observed trajectories. Second panel is the polynomial fit to the subsample. Third panel is the natural log fit to the subsample.

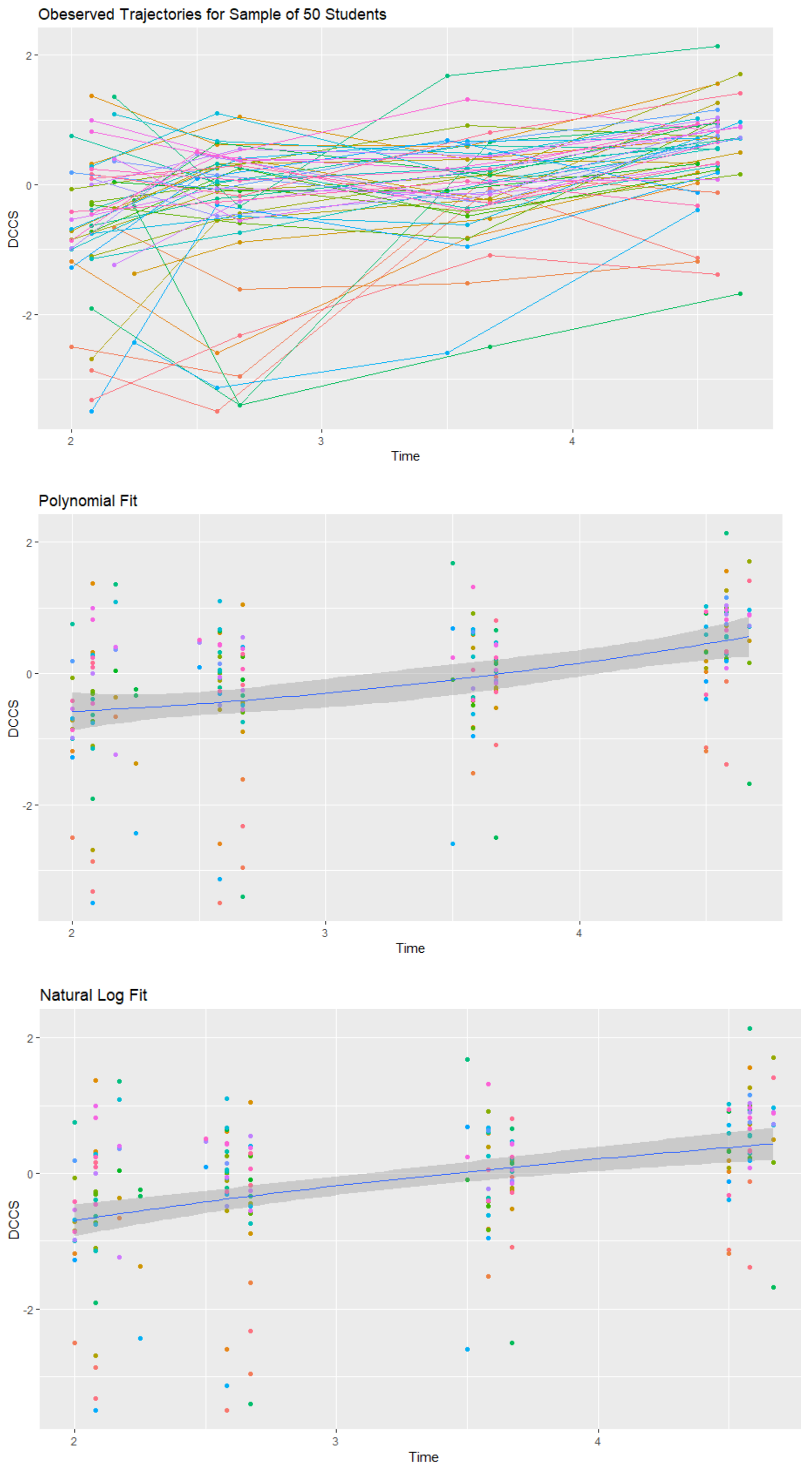

Figure 3 : Plots of cognitive flexibility trajectories for random subsample of 50 students. First panel is the observed trajectories. Second panel is the polynomial fit to the subsample. Third panel is the natural log fit to the subsample.

Research Question 1: *What is the average math achievement growth trajectory from kindergarten through fourth grade, accounting for the nesting of students within schools, and summer loss?*

The first step to answering research question 1 is to determine how to model the random effects. In theory, the intercept, the log trajectory, and summer drop may all vary at both the student and the school level. This is captured in the following model:

$$MA_{tis} = b_0 + b_1 ln(time)_{tis} + b_2 drop_{tis} + u_{0s} + u_{0is} + u_{1s} ln(time)_{tis} + u_{1is} ln(time)_{tis} + u_{2s} drop_{tis} + u_{2is} drop_{tis} + r_{tis}$$

The school random effects are summarized with an unstructured 3x3 matrix with variance components on the diagonal and covariances on the off-diagonals. The student-level random effects are similarly summarized and assumed to be independent of the school-level effects. However, overparameterized models – that is, those with too many random effects relative to the variability in the data – can produce near singular solutions in which the variance component estimates are near zero and/or correlations between random effects approach one. The extent to which this occurs varies across the three outcomes (math achievement, working memory, and cognitive flexibility). For this reason, each outcome will be summarized with four models. The first will be the full model presented above. Based on some evidence that the drop variable has minimal variability at the school level, a second model will be fit that removes this term.

$$MA_{tis} = b_0 + b_1 ln(time)_{tis} + b_2 drop_{tis} + u_{0s} + u_{0is} + u_{1s} ln(time)_{tis} + u_{1is} ln(time)_{tis} + u_{2is} drop_{tis} + r_{tis}$$

A third model will be presented removing drop from the student-level as well.

$$MA_{tis} = b_0 + b_1 ln(time)_{tis} + b_2 drop_{tis} + u_{0s} + u_{0is} + u_{1s} ln(time)_{tis} + u_{1is} ln(time)_{tis} + r_{tis}$$

Finally, a model with only a random coefficient for time at the student level will be fit.

$$MA_{tis} = b_0 + b_1 ln(time)_{tis} + b_2 drop_{tis} + u_{0s} + u_{0is} + u_{1is} ln(time)_{tis} + r_{tis}$$

The interpretation will focus where possible on the fully specified model with all random effects. However, the other three models are presented for the reader concerned about the size of the variance components and covariances (reported as correlations) in the tables. Information criteria (AIC and BIC) are also presented so that the reader can evaluate the degradation in model fit that occurs from constraining some of the variance components to be zero. The change in these statistics from one model to the next is also reported.

Table 3 presents the results from fitting growth models to the math achievement outcome that do not include any demographics. The first model represents the full specification with random intercepts, growth trajectories, and summer drop values estimated at the student and school levels. The time variable is scaled so that log(time) = 0 in kindergarten, and the math achievement outcome has been scaled to z-scores. The fixed effects results show a significant, positive trend (b = 1.751, p < 0.001) over time after accounting for a significant negative drop in fall semester (b = -0.051, p < 0.001) as well as school and student random effects.

The student-level random effects show that individuals with lower starting trajectories tend to have more positive trajectories, as evidenced by the negative correlation (r = -0.411). Likewise, those with lower than average starting points tend to have a more negative summer drop (r = 0.183). Students with higher than average trajectories were much more likely to have more negative summer drops (r = -0.856).

The school level random effects show that schools having lower than average starting values tended to have higher than average trajectories (r = -0.581). Schools with lower than average starting values also had more negative summer drops (r = 0.421). Summer drop and trajectories were negatively associated (r = -0.919), which indicates that those with higher than average trajectories had more negative summer drops.

The next part of the table presents the amount of variability in the intercept, trajectory, and summer drop that exists at the school level relative to the student level. Initial scores (the intercept) and trajectories were most variable at the student level. For the intercept, 22.8% of variability existed at the school level, while 29.9% of growth variability was attributable to schools. The variability in summer drop was equally distributed between school and student levels.

Table 3

Mixed Models of Math Achievement - No Demographics

	Model 1		Model 2		Model 3		Model 4	
Effect	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
Fixed Effects								
Intercept	-1.645***	0.011	-1.645***	0.0104	-1.644***	0.0104	-1.644***	0.0091
Time (Log)	1.751***	0.0094	1.751***	0.0063	1.749***	0.0057	1.748***	0.0048
Summer Drop	-0.051***	0.0031	-0.052***	0.0019	-0.051***	0.0017	-0.05***	0.0017
Random Effects								
Student - Var(Intercept)	0.261		0.266		0.253		0.264	
Student - Cor(Intercept, Log Time)	-0.411		-0.423		-0.457		-0.494	
Student - Cor(Intercept, Drop)	0.183		0.227		--		--	
Student - Var(Log Time)	0.11		0.144		0.041		0.048	
Student - Cor(Log Time, Drop)	-0.856		-0.875		--		--	
Student - Var(Drop)	0.005		0.01		--		--	
School - Var(Intercept)	0.077		0.066		0.067		0.047	
School - Cor(Intercept, Log Time)	-0.581		-0.557		-0.575		--	
School - Cor(Intercept, Drop)	0.421		--		--		--	
School - Var(Log Time)	0.047		0.007		0.008		--	
School - Cor(Log Time, Drop)	-0.919		--		--		--	
School - Var(Drop)	0.005		--		--		--	
Residual	0.048		0.048		0.05		0.05	
Percentage Variance at School Level								
Intercept	22.8%		19.9%		20.9%		15.1%	
Log Time	29.9%		4.6%		16.3%		--	
Summer Drop	50.0%		--		--		--	
Model Fit								
-2 Res Log Likelihood	48249		48881		49569		50320	
Δ -2 Res Log Likelihood	--		632		688		751	
AIC (Smaller is Better)	48275		48901		49583		50330	
Δ AIC	--		626		682		747	
BIC (Smaller is Better)	48336		48948		49616		50354	
Δ BIC	--		612		668		738	

Note. *** $p < 0.001$

Note that the correlation estimates between log time and summer drop were very high at both the student and school levels. This may accurately capture the strength of the association or may be due to estimates approaching the boundary of the parameter space given the relatively small summer drop variance component estimates (0.005 at both the student and school levels). Models 2 and 3 therefore incrementally remove the school-level and student-level summer drop variables. Because doing so reduces the size of the school-level trajectory random effect, Model 4 removes that term as well. However, the direction of the remaining random effects correlations are unchanged by these model alterations, and the fixed effects terms are the same across specifications. Thus, it is possible to have confidence in the overall inferences derived from the first model. To facilitate interpretation, Figure 4 displays the model-implied trajectory of math achievement for the average student.

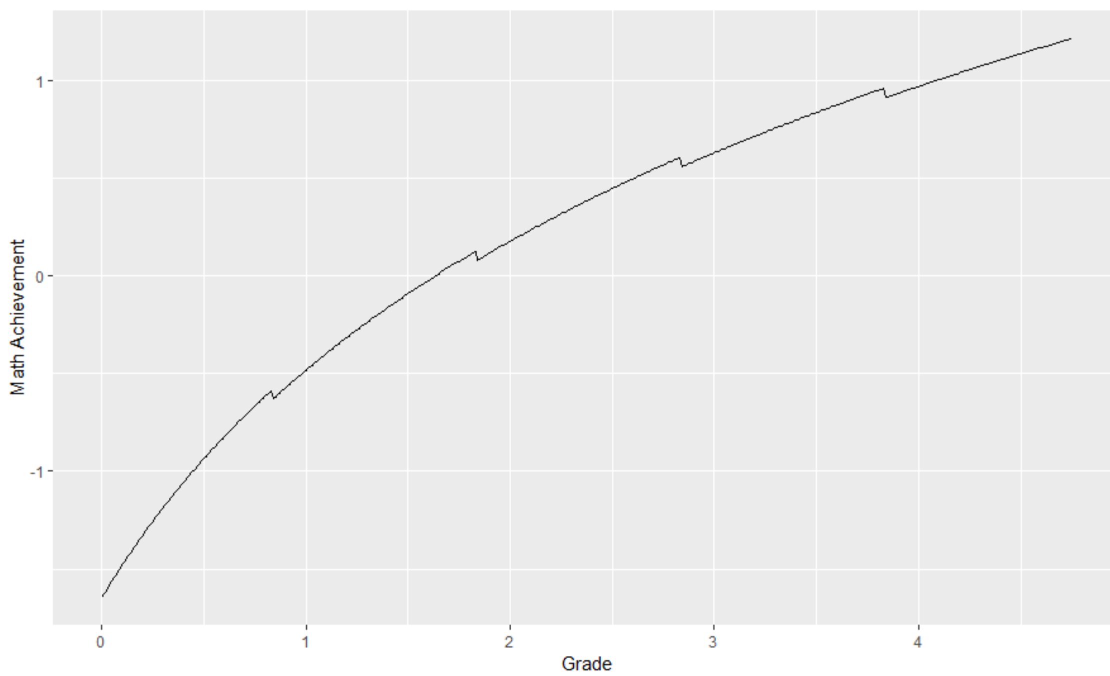

Figure 4: Model trajectory of math achievement based on model with no demographics.

Research Question 2: *To what degree do gender, racial, or socioeconomic status differences account for variation in the intercepts, slopes, and summer drops in the mathematics trajectories at the student and school levels?*

The primary model of interest is the full model from the prior section that also introduces demographics as follows.

$$\begin{aligned} MA_{tis} = b_0 &+ b_1 ln(time)_{tis} + b_2 drop_{tis} + b_3 gender_i + b_4 black_i + b_5 hispanic_i + b_6 asian_i \\ &+ b_7 other\ race_i + b_8 ses_i + b_9 gender_i ln(time)_{tis} + b_{10} black_i ln(time)_{tis} \\ &+ b_{11} hispanic_i ln(time)_{tis} + b_{12} asian_i ln(time)_{tis} \\ &+ b_{13} other\ race_i ln(time)_{tis} + b_{14} ses_i ln(time)_{tis} + b_{15} gender_i drop_{tis} \\ &+ b_{16} black_i drop_{tis} + b_{17} hispanic_i drop_{tis} + b_{18} asian_i drop_{tis} \\ &+ b_{19} other\ race_i drop_{tis} + b_{20} ses_i drop_{tis} + u_{0s} + u_{0is} + u_{1s} ln(time)_{tis} \\ &+ u_{1is} ln(time)_{tis} + u_{2s} drop_{tis} + u_{2is} drop_{tis} + r_{tis} \end{aligned}$$

Coefficients b_3 through b_8 represent the main effects of each demographics variable, or the effect when time = 0 (fall of kindergarten year). Coefficients b_9 through b_{14} are the interactions with time, which estimate how much the trajectories vary as a function of each demographics variable. Coefficients b_{15} through b_{20} are the interactions with summer drop, which estimate how much the effect of summer drop varies as a function of each demographics variable. The random effects u_{0s} through u_{2is} summarize how much the intercept, trajectory, and summer drop effects vary at the student and school levels net of any variability captured by the demographics. As was the case in the prior section, this model is the focus of interpretation. However, alternative models that incrementally constrain u_{2s}, u_{2is}, and u_{1s} to zero will also be presented as alternatives. Table 4 presents the results from the models. The results show a significant positive main effect for log time (b = 1.768, $p < 0.001$). The main effect for summer drop is negative (b = -0.045, $p < 0.001$). At the start of the series, scores tend to be lower for blacks (b = -0.161, $p < 0.001$) and Hispanics (b = -0.219, $p < 0.001$) but higher for females (b = 0.024, $p < 0.01$) and those with higher SES (b = 0.244, $p < 0.001$).

Table 4

Mixed Models of Math Achievement - With Demographics

Effect	Model 1 Estimate	Model 1 SE	Model 2 Estimate	Model 2 SE	Model 3 Estimate	Model 3 SE	Model 4 Estimate	Model 4 SE
Fixed Effects								
Intercept	-1.574***	0.0102	-1.573***	0.01	-1.572***	0.01	-1.573***	0.0091
Time (Log)	1.768***	0.0118	1.766***	0.0097	1.764***	0.0087	1.763***	0.0082
Summer Drop	-0.045***	0.0041	-0.045***	0.0032	-0.044***	0.0029	-0.044***	0.0029
Black	-0.161***	0.0173	-0.163***	0.0171	-0.162***	0.017	-0.138***	0.0162
Hispanic	-0.219***	0.0139	-0.218***	0.0138	-0.219***	0.0137	-0.219***	0.0132
Asian	0.008	0.0203	0.006	0.0201	0.007	0.0199	-0.002	0.0194
Other Race	-0.023	0.0213	-0.019	0.0213	-0.019	0.0211	-0.024	0.0211
Female	0.024**	0.0092	0.023*	0.0093	0.023*	0.0092	0.022*	0.0093
SES	0.244***	0.0067	0.248***	0.0067	0.248***	0.0066	0.263***	0.0065
Time (Log) X Black	-0.139***	0.0198	-0.13***	0.0176	-0.13***	0.0157	-0.148***	0.0152
Time (Log) X Hispanic	0.094***	0.0159	0.091***	0.0146	0.094***	0.0129	0.094***	0.0126
Time (Log) X Asian	0	0.0253	0.008	0.0236	0.004	0.0216	0.011	0.0212
Time (Log) X Other Race	-0.006	0.0245	-0.021	0.0245	-0.021	0.0216	-0.017	0.0216
Time (Log) X Female	-0.052***	0.0105	-0.05***	0.0108	-0.049***	0.0095	-0.049***	0.0095
Time (Log) X SES	-0.073***	0.0077	-0.089***	0.0073	-0.089***	0.0065	-0.101***	0.0064
Summer Drop X Black	0.014*	0.0068	0.011	0.0059	0.011*	0.0054	0.011*	0.0054
Summer Drop X Hispanic	-0.023***	0.0055	-0.022***	0.0049	-0.023***	0.0045	-0.023***	0.0045
Summer Drop X Asian	0.026**	0.0089	0.023**	0.0083	0.024**	0.0078	0.024**	0.0078
Summer Drop X Other Race	0.003	0.0085	0.009	0.0085	0.009	0.0078	0.009	0.0078
Summer Drop X Female	-0.006	0.0036	-0.006	0.0037	-0.007	0.0034	-0.007	0.0034
Summer Drop X SES	0.02***	0.0027	0.026***	0.0025	0.026***	0.0023	0.026***	0.0023

Table 4 (*continued*)

Effect	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
Random Effects								
Student - Var(Intercept)	0.235		0.237		0.229		0.238	
Student - Cor(Intercept, Log Time)	-0.402		-0.399		-0.473		-0.504	
Student - Cor(Intercept, Drop)	0.143		0.172		--		--	
Student - Var(Log Time)	0.105		0.136		0.039		0.044	
Student - Cor(Log Time, Drop)	-0.859		-0.875		--		--	
Student - Var(Drop)	0.005		0.009		--		--	
School - Var(Intercept)	0.026		0.024		0.024		0.012	
School - Cor(Intercept, Log Time)	-0.548		-0.752		-0.729		--	
School - Cor(Intercept, Drop)	0.298		--		--		--	
School - Var(Log Time)	0.039		0.006		0.006		--	
School - Cor(Log Time, Drop)	-0.927		--		--		--	
School - Var(Drop)	0.004		--		--		--	
Residual	0.048		0.048		0.05		0.05	
Percentage Variance at School Level								
Intercept	10.0%		9.2%		9.5%		4.8%	
Log Time	27.1%		4.2%		13.3%		--	
Summer Drop	44.4%		--		--		--	
Model Fit								
-2 Res Log Likelihood	45390		45923		46623		47098	
Δ -2 Res Log Likelihood	--		533		700		475	
AIC (Smaller is Better)	45416		45943		46637		47108	
Δ AIC	--		527		694		471	
BIC (Smaller is Better)	45477		45990		46669		47132	
Δ BIC	--		513		679		463	

Note. *** $p < 0.001$. ** $p < 0.01$. * $p < 0.05$.

The positive trajectory observed for the main effect tends to be less positive for black students ($b = -0.139$, $p < 0.001$), slightly more positive for Hispanics ($b = 0.094$, $p < 0.001$), slightly less positive for females ($b = -0.052$, $p < 0.001$), and less positive as SES increases ($b = 0.073$, $p < 0.001$). The negative main effect for summer drop tends to be more negative for Hispanics ($b = -0.023$, $p < 0.001$) but less negative for blacks ($b = 0.014$, $p < 0.05$), Asians ($b = 0.026$, $p < 0.01$), and those with higher SES ($b = 0.02$, $p < 0.001$).

Turning to the random effects, net of demographics effects, students who start out higher tend to have lower trajectories ($r = -0.402$) and less negative summer drop effects ($r = 0.143$). The association between the trajectory and summer drop is large and negative ($r = -0.859$). At the school level, schools with higher starting levels tend to have less steep trajectories ($r = -0.548$) and less negative summer drop effects ($r = 0.298$). The correlation between summer drop and trajectories is very strongly negative ($r = -0.927$).

Even after controlling for demographics, the bulk of variability in intercepts, trajectories, and summer drop effects are mostly at the student level. 10% of variability in starting scores (for white males with SES = 0) is attributable to school effects. 27.1% of trajectories is associated with school effects. A little less than half of the summer drop, or 44.4%, can be attributed to schools.

Table 5 summarizes how much the variance components are reduced after adding in the demographics variables. The first numeric column is the variance component estimates from the models without demographics, and the next column is the variance components from the models with demographics. The first panel of the table corresponds to Model 1, or the model with all variance components and covariances estimated. The demographics have much more of an impact on the school-level variances relative to the student-level variances. Demographics fixed effects reduce the student-level intercept variance by 8.11%, while the school-to-school intercept variability is reduced by 66.23%. For time, the student-level variance is reduced by 4.55% while the school variance is reduced by 17.02%. Finally, the summer drop variance at the student level is unaffected by demographics, while the school-level variance is reduced by 20%.

Table 5

Variance Explained - Math Achievement

Model	No Demographics	With Demographics	% Explained
Model 1			
Student Level			
Student - Var(Intercept)	0.261	0.235	8.11
Student - Var(Log Time)	0.11	0.105	4.55
Student - Var(Drop)	0.005	0.005	0.00
School Level			
School - Var(Intercept)	0.077	0.026	66.23
School - Var(Log Time)	0.047	0.039	17.02
School - Var(Drop)	0.005	0.004	20.00
Model 2			
Student - Var(Intercept)	0.27	0.24	10.90
Student - Var(Log Time)	0.14	0.14	5.56
Student - Var(Drop)	0.01	0.01	10.00
School Level			
School - Var(Intercept)	0.07	0.02	63.64
School - Var(Log Time)	0.01	0.01	14.29
Model 3			
Student Level			
Student - Var(Intercept)	0.253	0.229	9.49
Student - Var(Log Time)	0.041	0.039	4.88
School Level			
School - Var(Intercept)	0.067	0.024	64.18
School - Var(Log Time)	0.008	0.006	25.00
Model 4			
Student Level			
Student - Var(Intercept)	0.264	0.238	9.85
Student - Var(Log Time)	0.048	0.044	8.33
School Level			
School - Var(Intercept)	0.047	0.012	74.47

Changing the number of random effects estimated does little to impact the estimates of the fixed effects. Figure 5 presents a summary of the trajectories estimated from the fixed effects.

Because the intercept, trajectories, and summer drop all vary according to gender, race, and SES, the figure is faceted according to the demographics variables to provide a sense of the substantive effect of each interaction.

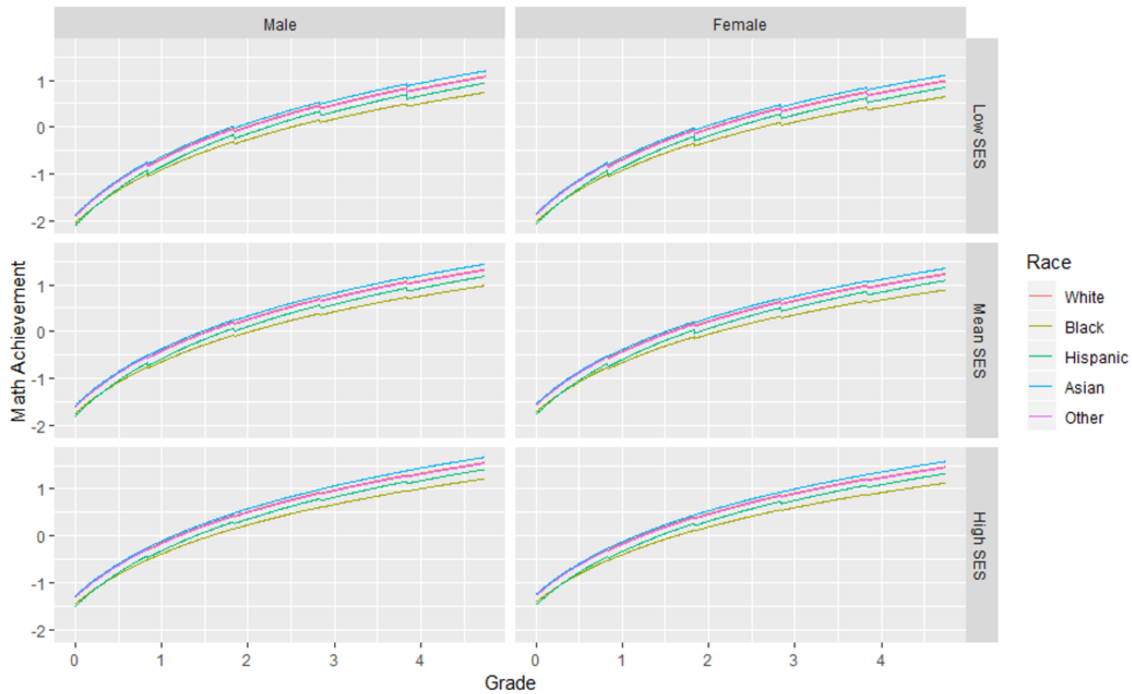

Figure 5 : Model trajectories of math achievement by race, gender, and SES. Low SES defined as 1.5 standard deviations below the mean. High SES defined as 1.5 standard deviations above the mean.

Research Question 3: *Is executive functioning associated with change in math achievement over this period?*

The third research question requires first coming up with specifications for the two executive functioning models. The working memory model will be determined following the same steps as were used for math achievement. First, a model with all possible random effects is specified.

$$WM_{tis} = b_0 + b_1 ln(time)_{tis} + b_2 drop_{tis} + u_{0s} + u_{0is} + u_{1s} ln(time)_{tis} + u_{1is} ln(time)_{tis} + u_{2s} drop_{tis} + u_{2is} drop_{tis} + r_{tis}$$

Additional models will then be fit to account for concerns that variance components and their correlations are too close to the boundary values. These three models are the following:

$$WM_{tis} = b_0 + b_1 ln(time)_{tis} + b_2 drop_{tis} + u_{0s} + u_{0is} + u_{1s} ln(time)_{tis} + u_{1is} ln(time)_{tis} + u_{2is} drop_{tis} + r_{tis}$$

$$WM_{tis} = b_0 + b_1 ln(time)_{tis} + b_2 drop_{tis} + u_{0s} + u_{0is} + u_{1s} ln(time)_{tis} + u_{1is} ln(time)_{tis} + r_{tis}$$

$$WM_{tis} = b_0 + b_1 ln(time)_{tis} + b_2 drop_{tis} + u_{0s} + u_{0is} + u_{1is} ln(time)_{tis} + r_{tis}$$

The cognitive flexibility outcome will be modeled slightly differently due to the fact that the DCCS instrument was not administered until the second grade. Specifically, the drop variable will be removed from both fixed and random effects specifications. The primary model will be

$$CF_{tis} = b_0 + b_1 ln(time)_{tis} + u_{0s} + u_{0is} + u_{1s} ln(time)_{tis} + u_{1is} ln(time)_{tis} + r_{tis}$$

In the presence of small school-level effects on trajectories, the following model will also be fit:

$$CF_{tis} = b_0 + b_1 ln(time)_{tis} + u_{0s} + u_{0is} + u_{1s} ln(time)_{tis} + r_{tis}$$

Table 6 presents the results for the working memory models. Based on the first model (but consistent across random effects specifications), the overall trajectory is significant and positive (b = 1.311, $p < 0.001$). The effect of summer drop is also significant and tends to reduce scores (b = -0.06, $p < 0.001$).

Table 6

Mixed Models of Working Memory - No Demographics

Effect	Model 1 Estimate	SE	Model 2 Estimate	SE	Model 3 Estimate	SE	Model 4 Estimate	SE
Fixed Effects								
Intercept	-1.196***	0.0147	-1.196***	0.0144	-1.196***	0.0143	-1.197***	0.0106
Time (Log)	1.311***	0.0153	1.312***	0.0132	1.31***	0.0122	1.31***	0.0111
Summer Drop	-0.06***	0.0052	-0.06***	0.0044	-0.059***	0.0041	-0.059***	0.0041
Random Effects								
Student - Var(Intercept)	0.53		0.535		0.508		0.552	
Student - Cor(Intercept, Log Time)	-0.549		-0.552		-0.732		-0.759	
Student - Cor(Intercept, Drop)	0.182		0.206		--		--	
Student - Var(Log Time)	0.46		0.501		0.129		0.15	
Student - Cor(Log Time, Drop)	-0.867		-0.878		--		--	
Student - Var(Drop)	0.033		0.039		--		--	
School - Var(Intercept)	0.12		0.112		0.112		0.038	
School - Cor(Intercept, Log Time)	-0.691		-0.901		-0.896		--	
School - Cor(Intercept, Drop)	0.26		--		--		--	
School - Var(Log Time)	0.068		0.02		0.021		--	
School - Cor(Log Time, Drop)	-0.84		--		--		--	
School - Var(Drop)	0.006		--		--		--	
Residual	0.281		0.281		0.289		0.289	
Percentage Variance at School Level								
Intercept	18.5%		17.3%		18.1%		6.4%	
Log Time	12.9%		3.8%		14.0%		0.0%	
Summer Drop	15.4%		0.0%		--		--	
Model Fit								
-2 Res Log Likelihood	221343		221393		221898		222399	
Δ -2 Res Log Likelihood	--		50		505		501	
AIC (Smaller is Better)	221369		221413		221912		222409	
Δ AIC	--		44		499		497	
BIC (Smaller is Better)	221430		221460		221945		222432	
Δ BIC	--		30		485		487	

Note. *** $p < 0.001$

The random effects show that higher starting scores tend to be associated with lower trajectories at the student level (r = -0.549), while higher starting scores are associated with a less negative effect of summer drop (r = 0.182). The trajectory and drop variance components are strongly negatively associated (r = -0.867), indicating that those with higher starting values exhibit greater negative drop effects. A similar pattern happens at the school level. Schools with higher scores in fall of kindergarten show weaker trajectories (r = -0.691) as well as the least negative effects of summer (r = 0.26). The association between trajectories and summer drop is strongly negative (r = -0.84). Figure 6 presents a visualization of the modeled trajectory.

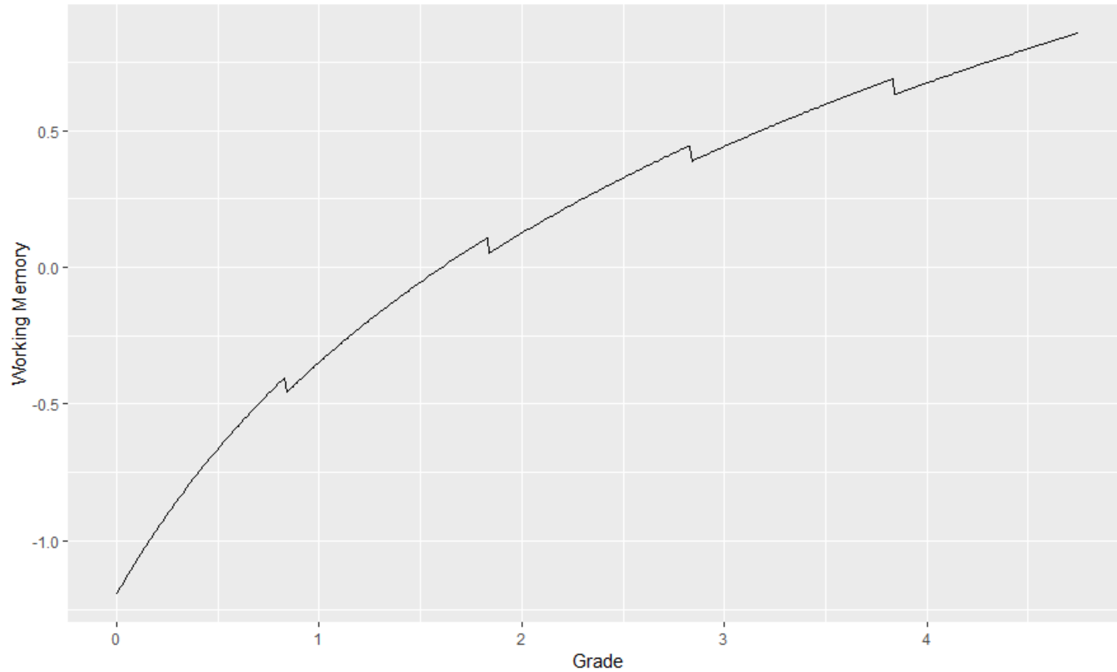

Figure 6: Model trajectory of working memory based on model with no demographics.

The bulk of the variability in working memory scores, trajectories, and summer drop effects occur at the student level. 18.5% of the variability in intercepts is due to school-level factors. 12.9% of the trajectory variance is associated with school effects. 15.4% of summer drop effects are associated with schools.

To specifically address the research question, the random effects from the math achievement model were correlated with the random effects of the working memory model. This was performed by fitting each model, estimating the best linear unbiased predictors (BLUPs) of the student- and school-level random effects, and then correlating them. Table 7 presents results at the student level, and Table 8 presents results at the school level.

Model 1 in Table 7 corresponds to the model with all possible random effects included and is the focus of interpretation. Given the very high correlations between trajectories and summer drop within the math achievement and working memory models already described, the correlations for the remaining three models are also presented for reference. The results show that higher starting scores on working memory tend to be associated with weaker math achievement trajectories (*r* = -0.105). Equivalently, lower starting scores in working memory tend to be associated with greater improvements in math achievement. In addition, the correlations show that higher scores in kindergarten on working memory tend to be strongly associated with higher scores in kindergarten on math achievement (*r* = 0.711). Higher starting scores on working memory tend to be associated with more negative math achievement summer drops (*r* = -0.081). Higher working memory trajectories are associated with higher math achievement trajectories (*r* = 0.238) but more negative summer drop effects (*r* = -0.200). Less negative summer drop effects for working memory tend to be associated with more negative math achievement intercepts (*r* = -0.25), weaker math achievement trajectories (*r* = -0.185), and less negative math achievement summer drops (*r* = 0.257).

Table 7

Correlations between Math Achievement and Working Memory Student-Level Random Effects - Unadjusted Models

	MA Intercept	MA Log Time	MA Drop	EF Intercept	EF Log Time	EF Drop
Model 1						
MA Intercept	1					
MA Log Time	-0.332	1				
MA Drop	0.069	-0.929	1			
EF Intercept	0.711	-0.105	-0.081	1		
EF Log Time	-0.146	0.238	-0.2	-0.444	1	
EF Drop	-0.25	-0.185	0.257	-0.101	-0.825	1
Model 2						
MA Intercept	1					
MA Log Time	-0.346	1				
MA Drop	0.133	-0.904	1			
EF Intercept	0.708	-0.109	-0.031	1		
EF Log Time	-0.145	0.237	-0.186	-0.444	1	
EF Drop	-0.222	-0.188	0.22	-0.063	-0.841	1
Model 3						
MA Intercept	1					
MA Log Time	-0.406	1				
EF Intercept	0.738	-0.184		1		
EF Log Time	-0.46	0.273		-0.736	1	
Model 4						
MA Intercept	1					
MA Log Time	-0.458	1				
EF Intercept	0.726	-0.21		1		
EF Log Time	-0.47	0.292		-0.773	1	

Note. All correlations significant at $p < 0.001$.

Table 8

Correlations between Math Achievement and Working Memory School-Level Random Effects - Unadjusted Models

	MA Intercept	MA Log Time	MA Drop	EF Intercept	EF Log Time	EF Drop
Model 1						
MA Intercept	1					
MA Log Time	-0.609	1				
MA Drop	0.454	-0.923	1			
EF Intercept	0.849	-0.374	0.266	1		
EF Log Time	-0.601	0.457	-0.379	-0.778	1	
EF Drop	0.124	-0.33	0.339	0.241	-0.763	1
Model 2						
MA Intercept	1					
MA Log Time	-0.577	1				
EF Intercept	0.87	-0.374		1		
EF Log Time	-0.798	0.422		-0.945	1	
Model 3						
MA Intercept	1					
MA Log Time	-0.597	1				
EF Intercept	0.87	-0.393		1		
EF Log Time	-0.793	0.443		-0.94	1	
Model 4						
MA Intercept	1					
EF Intercept	0.852			1		

Note. All correlations significant at $p< 0.001$.

Table 8 presents the same information for school-level random effects. The story is very similar. Schools with higher working memory intercepts are associated with having lower math achievement trajectories (*r* = -0.374). The two random intercepts are strongly, positively correlated (*r* = 0.849), indicating that those whose working memory scores are high at baseline also tend to have higher math achievement. Higher working memory intercepts are also associated with less negative summer drops (*r* = 0.266). The working memory trajectory is positively associated with the math achievement trajectory (*r* = 0.457) but negatively associated with math achievement summer drop (*r* = -0.379). Executive functioning summer drop tends to be less negative when math achievement baselines are higher (*r* = 0.124) but more negative when math achievement trajectories are higher (*r* = -0.33). Both summer drop effects are positively associated (*r* = 0.339).

Table 9 turns to presenting results for modeling cognitive flexibility. For this model, time has been rescaled to make zero = second grade. Thus, the intercept fixed effect is the estimated average cognitive flexibility score at the start of the available series. The log time fixed effect is positive and highly significant ($b = 0.777$, $p < 0.001$). At the student level, there is a large, negative association between the random intercept and random trajectory (*r* = -0.873). This is repeated at the school level (*r* = -0.954). Only 11.4% of the variability in baseline values is attributable to schools. Only 4.9% of variability in trajectories is attributable to schools. Figure 7 presents a visualization of the trajectory.

Table 10 presents the correlations in student-level random effects between math achievement and cognitive flexibility. The association between the cognitive flexibility intercept and math achievement trajectory is negative and significant (*r* = -0.255). Lower levels of cognitive flexibility in kindergarten mean that math achievement will improve more over time. The two intercepts are strongly correlated (*r* = 0.522), indicating that higher cognitive flexibility scores among students at baseline go along with higher baseline math achievement. The math achievement intercept and executive functioning trajectories are negatively associated (r = -0.469), and both trajectories are positively related (r = 0.230).

Table 9
Mixed Models of Cognitive Flexibility - No Demographics

Effect	Model 1 Estimate	SE	Model 2 Estimate	SE
Fixed Effects				
Intercept	-0.539***	0.0148	-0.54***	0.0118
Time (Log)	0.777***	0.0086	0.778***	0.0074
Random Effects				
Student - Var(Intercept)	0.824		0.856	
Student - Cor(Intercept, Log Time)	-0.873		-0.881	
Student - Var(Log Time)	0.313		0.328	
School - Var(Intercept)	0.106		0.044	
School - Cor(Intercept, Log Time)	-0.954		--	
School - Var(Log Time)	0.016		--	
Residual	0.371		0.371	
Percentage Variance at School Level				
Intercept	11.4%		4.9%	
Log Time	4.9%		--	
Model Fit				
-2 Res Log Likelihood	133647		133801	
Δ -2 Res Log Likelihood	--		154	
AIC (Smaller is Better)	133661		133811	
Δ AIC	--		150	
BIC (Smaller is Better)	133694		133835	
Δ BIC	--		141	

Note. *** $p < 0.001$

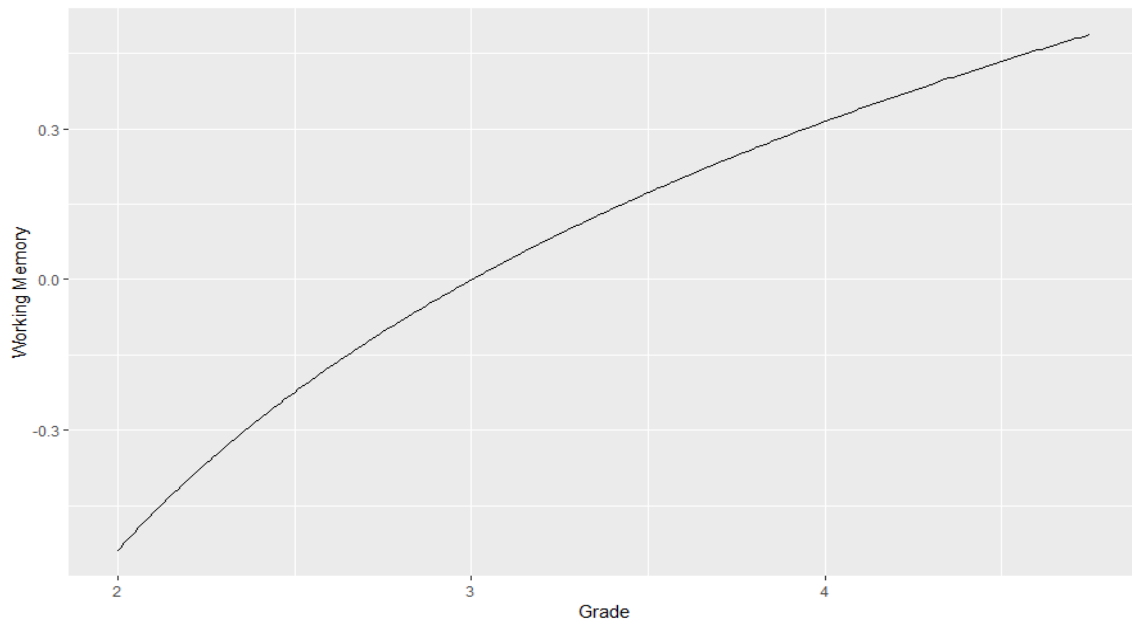

Figure 7: Model trajectory of Cognitive Flexibility based on model with no demographics. In the statistical model, time was rescaled so that grade 2 = zero.

Table 10

Correlations between Math Achievement and Cognitive Flexibility

Student-Level Random Effects - Unadjusted Models

	MA Intercept	MA Log Time	EF Intercept	EF Log Time
Model 1				
MA Intercept	1			
MA Log Time	-0.508	1		
EF Intercept	0.522	-0.255	1	
EF Log Time	-0.469	0.230	-0.940	1
Model 2				
MA Intercept	1			
MA Log Time	-0.517	1		
EF Intercept	0.548	-0.248	1	
EF Log Time	-0.470	0.226	-0.946	1

Note. All correlations significant at $p < 0.05$.

A similar story occurs for the school-level random effects, as shown in Table 11. The cognitive flexibility intercept and math achievement trajectory are negatively associated (*r* = -0.206). cognitive flexibility and math achievement intercepts are strongly, positively correlated (*r* = 0.710).Higher cognitive flexibility trajectories are associated with lower math achievement intercepts (*r* = -0.713) and but positively associated with math achievement trajectories (*r* = 0.212).

Table 11

Correlations between Math Achievement and Cognitive Flexibility School-Level Random Effects - Unadjusted Models

	MA Intercept	MA Log Time	EF Intercept	EF Log Time
Model 1				
MA Intercept	1			
MA Log Time	-0.374	1		
EF Intercept	0.710	-0.206	1	
EF Log Time	-0.713	0.212	-0.995	1
Model 2				
MA Intercept	1			
EF Intercept	0.668		1	

Note. All correlations significant at $p < 0.05$.

The very high correlation between the cognitive flexibility intercept and trajectory at both the student and school levels suggest that pathologies in the numeric optimization occurred due to estimates at the parameter space boundary. For this reason, a second model is provided that constrains the school-level random trajectory to be zero. However, the same issue occurs at the student level. Constraining the student-level random trajectory to zero would not allow for any inferences about how the math achievement and cognitive flexibility trends go together, so no further model is provided. However, it is advised that the results be taken with the caveat that the estimates may not be fully reliable.

Research Question 4: *Which student sociodemographic characteristics are most strongly associated change in math achievement and to what degree does controlling for student sociodemographic characteristics alter the association between executive functioning and change in math achievement?*

The final research question is similar to the last except that the models now incorporate demographics. For working memory, the model of interest is the following:

$$
\begin{aligned}
WM_{tis} = b_0 &+ b_1 ln(time)_{tis} + b_2 drop_{tis} + b_3 gender_i + b_4 black_i + b_5 hispanic_i + b_6 asian_i \\
&+ b_7 other\ race_i + b_8 ses_i + b_9 gender_i ln(time)_{tis} + b_{10} black_i ln(time)_{tis} \\
&+ b_{11} hispanic_i ln(time)_{tis} + b_{12} asian_i ln(time)_{tis} \\
&+ b_{13} other\ race_i ln(time)_{tis} + b_{14} ses_i ln(time)_{tis} + b_{15} gender_i drop_{tis} \\
&+ b_{16} black_i drop_{tis} + b_{17} hispanic_i drop_{tis} + b_{18} asian_i drop_{tis} \\
&+ b_{19} other\ race_i drop_{tis} + b_{20} ses_i drop_{tis} + u_{0s} + u_{0is} + u_{1s} ln(time)_{tis} \\
&+ u_{1is} ln(time)_{tis} + u_{2s} drop_{tis} + u_{2is} drop_{tis} + r_{tis}
\end{aligned}
$$

As was the case for math achievement, additional models with fewer random effects will also be presented. The model for cognitive flexibility is similar but removes summer drop from both fixed and random effects.

$$
\begin{aligned}
CF_{tis} = b_0 &+ b_1 ln(time)_{tis} + b_2 gender_i + b_3 black_i + b_4 hispanic_i + b_5 asian_i \\
&+ b_6 other\ race_i + b_7 ses_i + b_8 gender_i ln(time)_{tis} + b_9 black_i ln(time)_{tis} \\
&+ b_{10} hispanic_i ln(time)_{tis} + b_{11} asian_i ln(time)_{tis} \\
&+ b_{12} other\ race_i ln(time)_{tis} + b_{13} ses_i ln(time)_{tis} + u_{0s} + u_{0is} \\
&+ u_{1s} ln(time)_{tis} + u_{1is} ln(time)_{tis} + r_{tis}
\end{aligned}
$$

The model results for working memory are presented in Table 12. All of the demographics variables play a role in the fixed effects. The main effect of log time is positive and significant ($b = 1.255$, $p < 0.001$), which corresponds to the expected trajectory for a white male with the SES scale equal to zero (a little above the mean). The main effect for summer drop is negative and significant ($b = -0.062$, $p < 0.001$). The demographics main effects show that working memory scores in kindergarten tend to be significantly lower for blacks ($b = -0.294$, $p < 0.001$), Hispanics ($b = -0.306$, $p < 0.001$), and slightly lower for Asians ($b = -0.071$, $p < 0.05$). Females tend to have higher scores in kindergarten ($b = 0.074$, $p < 0.001$), as do those with higher SES ($b = 0.314$, $p < 0.001$). Hispanic trajectories tend to be more positive relative to white students ($b = 0.138$, $p < 0.001$), while higher SES tends to reduce the strength of the trajectories ($b = -0.169$, $p < 0.001$). The negative effect of summer drop is somewhat less for blacks ($b = 0.044$, $p < 0.01$), those of other races ($b = 0.044$, $p < 0.05$), and those with higher SES ($b = 0.03$, $p < 0.001$). The effect is more strongly negative for females ($b = -0.019$, $p < 0.05$). Figure 8 presents a graph summarizing the trajectories based on the fixed effects.

Turning to the random effects, students who begin with higher intercepts tend to have less positive trajectories ($r = -0.54$) but less negative summer drop effects ($r = 0.178$). The trajectory and summer drop effects are negatively associated ($r = -0.870$). These patterns also occur at the school level. Schools with higher starting values in kindergarten tend to have lower trajectories ($r = -0.581$) and less negative summer drop effects ($r = 0.302$). Summer drop and trajectories are negatively related ($r = -0.918$).

Table 13 shows the amount of variability between students and schools that is explained by the inclusion of random effects. As was the case for math achievement, demographics fixed effects reduce the school-level variance components more than the student-level variance components. Demographics reduce the student-level random intercept by 8.11%, while the school-level random intercept is reduced by 72.5%. The trajectory variance is reduced by 2.39% for students and by 26.47% for schools. Neither the student-level nor the school-level summer drop variance is affected by demographics.

Table 12

Mixed Models of Working Memory - With Demographics

	Model 1		Model 2		Model 3		Model 4	
Effect	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
Fixed Effects								
Intercept	-1.107***	0.0151	-1.108***	0.0149	-1.107***	0.0148	-1.103***	0.0139
Time (Log)	1.255***	0.0226	1.258***	0.0211	1.256***	0.0193	1.254***	0.019
Summer Drop	-0.062***	0.0081	-0.063***	0.0075	-0.063***	0.007	-0.063***	0.007
Black	-0.294***	0.0271	-0.293***	0.0268	-0.293***	0.0266	-0.292***	0.0254
Hispanic	-0.306***	0.0221	-0.303***	0.0219	-0.303***	0.0217	-0.314***	0.021
Asian	-0.071*	0.0339	-0.074*	0.0336	-0.073*	0.0333	-0.081*	0.0324
Other Race	0.037	0.0353	0.039	0.0352	0.041	0.0349	0.037	0.0347
Female	0.074***	0.0154	0.073***	0.0154	0.072***	0.0152	0.072***	0.0153
SES	0.314***	0.0109	0.315***	0.0108	0.314***	0.0107	0.326***	0.0104
Time (Log) X Black	-0.012	0.0412	-0.015	0.0389	-0.015	0.0356	-0.017	0.0352
Time (Log) X Hispanic	0.138***	0.0338	0.129***	0.0323	0.13***	0.0295	0.136***	0.0292
Time (Log) X Asian	0.104	0.0554	0.11*	0.0535	0.108*	0.05	0.113*	0.0497
Time (Log) X Other Race	-0.102	0.0554	-0.107	0.0549	-0.114*	0.0502	-0.111*	0.0501
Time (Log) X Female	0.031	0.0241	0.032	0.0242	0.034	0.0221	0.035	0.0221
Time (Log) X SES	-0.169***	0.0169	-0.172***	0.0164	-0.171***	0.015	-0.179***	0.0149
Summer Drop X Black	0.044**	0.0148	0.045**	0.014	0.045***	0.0131	0.046***	0.0131
Summer Drop X Hispanic	0.019	0.0122	0.022	0.0116	0.022*	0.0109	0.022*	0.0109
Summer Drop X Asian	0.019	0.0203	0.016	0.0196	0.017	0.0186	0.017	0.0186
Summer Drop X Other Race	0.044*	0.0201	0.046*	0.02	0.049**	0.0187	0.049**	0.0187
Summer Drop X Female	-0.019*	0.0088	-0.02*	0.0088	-0.021*	0.0082	-0.021*	0.0082
Summer Drop X SES	0.03***	0.0061	0.031***	0.0059	0.031***	0.0055	0.031***	0.0055

Table 12 (*continued*)

	Model 1		Model 2		Model 3		Model 4	
Effect	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
Random Effects								
Student - Var(Intercept)	0.487		0.49		0.468		0.482	
Student - Cor(Intercept, Log Time)	-0.54		-0.54		-0.722		-0.732	
Student - Cor(Intercept, Drop)	0.178		0.195		--		--	
Student - Var(Log Time)	0.449		0.487		0.124		0.132	
Student - Cor(Log Time, Drop)	-0.87		-0.88		--		--	
Student - Var(Drop)	0.033		0.038		--		--	
School - Var(Intercept)	0.033		0.029		0.029		0.012	
School - Cor(Intercept, Log Time)	-0.581		-0.791		-0.775		--	
School - Cor(Intercept, Drop)	0.302		--		--		--	
School - Var(Log Time)	0.05		0.008		0.008		--	
School - Cor(Log Time, Drop)	-0.918		--		--		--	
School - Var(Drop)	0.006		--		--		--	
Residual	0.281		0.281		0.289		0.289	
Percentage Variance at School Level								
Intercept	6.3%		5.6%		5.8%		2.4%	
Log Time	10.0%		1.6%		6.1%		--	
Summer Drop	15.4%		--		--		--	
Model Fit								
-2 Res Log Likelihood	219666		219714		220231		220314	
Δ -2 Res Log Likelihood	--		48		517		83	
AIC (Smaller is Better)	219692		219734		220245		220324	
Δ AIC	--		42		511		79	
BIC (Smaller is Better)	219753		219781		220277		220347	
Δ BIC	--		28		496		70	

Note. *** $p < 0.001$. ** $p < 0.01$. * $p < 0.05$.

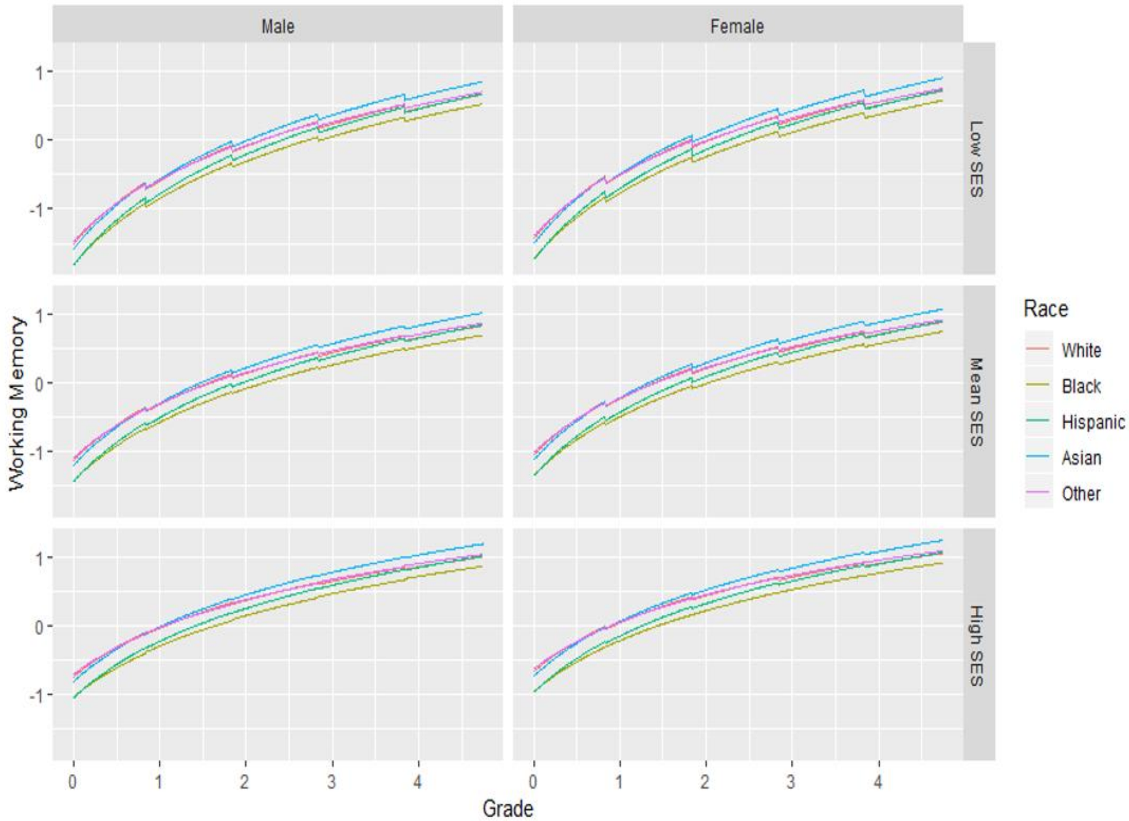

Figure 8: Model trajectories of working memory by race, gender, and SES. Low SES defined as 1.5 standard deviations below the mean. High SES defined as 1.5 standard deviations above the mean.

Table 13

Variance Explained - Working Memory

Model	No Demographics	With Demographics	% Explained
Model 1			
Student Level			
Student - Var(Intercept)	0.53	0.487	8.113
Student - Var(Log Time)	0.46	0.449	2.391
Student - Var(Drop)	0.033	0.033	0.000
School Level			
School - Var(Intercept)	0.12	0.033	72.500
School - Var(Log Time)	0.068	0.05	26.471
School - Var(Drop)	0.006	0.006	0.000
Model 2			
Student Level			
Student - Var(Intercept)	0.535	0.49	8.411
Student - Var(Log Time)	0.501	0.487	2.794
Student - Var(Drop)	0.039	0.038	2.564
School Level			
School - Var(Intercept)	0.112	0.029	74.107
School - Var(Log Time)	0.02	0.008	60.000
Model 3			
Student Level			
Student - Var(Intercept)	0.508	0.468	7.874
Student - Var(Log Time)	0.129	0.124	3.876
School Level			
School - Var(Intercept)	0.112	0.029	74.107
School - Var(Log Time)	0.021	0.008	61.905
Model 4			
Student Level			
Student - Var(Intercept)	0.552	0.482	12.681
Student - Var(Log Time)	0.15	0.132	12.000
School Level			
School - Var(Intercept)	0.038	0.012	68.421

Table 14 presents the correlations between math achievement and working memory random effects at the student level. Model 1 represents the model that includes all possible random effects, the remaining models are presented for reference. Based on the Model 1 results, the working memory intercept has a negligible association with the math achievement time trajectory (*r* = -0.076). That is, after controlling for demographics, working memory in kindergarten is a weak predictor of the subsequent growth in math achievement. The working memory intercept random effects are highly correlated with the math achievement intercepts (*r* = 0.688) such that higher baseline scores on one tend to appear with higher scores on the other. The working memory intercept has a small negative association with the math achievement summer drop effect (*r* = -0.15). The working memory trajectory is negatively associated with math achievement intercepts (*r* = -0.106) and math achievement summer drop (*r* = -0.202) but is positively related to math achievement's trajectory (*r* = 0.236). Summer drop in working memory is negatively associated with math achievement's intercept (*r* = -0.273) and the math achievement trajectory (*r* = -0.197) but positively associated with the math achievement summer drop (*r* = 0.292).

Table 15 presents the results for the school-level random effects, where the patterns are similar to Table 12. The working memory intercept is negatively associated with the math achievement trajectory (r = -0.178), indicating that school-level working memory scores in kindergarten tend to go with weaker math trajectories. The working memory intercept is highly, positively correlated with the math achievement intercept (r = 0.622); there is little association with math achievement summer drop (r = 0.029). The working memory trajectory is negatively associated with the math achievement intercept (r = -0.202) and math achievement summer drop (r = -0.296) but positively associated with the math achievement trajectory (r = 0.331). Executive functioning summer drop has a scant relationship with the math achievement intercept (r = -0.023), a stronger negative association with math achievement trajectories (r = -0.296), and a medium-sized association with the math achievement summer drop random effect (r = 0.33).

Table 14

Correlations between Math Achievement and Working Memory Student-Level Random Effects - Adjusted Models

	MA Intercept	MA Log Time	MA Drop	EF Intercept	EF Log Time	EF Drop
Model 1						
MA Intercept	1					
MA Log Time	-0.306	1				
MA Drop	-0.022	-0.919	1			
EF Intercept	0.688	-0.076	-0.15	1		
EF Log Time	-0.106	0.236	-0.202	-0.417	1	
EF Drop	-0.273	-0.197	0.292	-0.123	-0.829	1
Model 2						
MA Intercept	1					
MA Log Time	-0.296	1				
MA Drop	0.012*ns*	-0.901	1			
EF Intercept	0.684	-0.062	-0.124	1		
EF Log Time	-0.098	0.233	-0.187	-0.411	1	
EF Drop	-0.256	-0.205	0.263	-0.098	-0.843	1
Model 3						
MA Intercept	1					
MA Log Time	-0.419	1				
EF Intercept	0.718	-0.183		1		
EF Log Time	-0.426	0.272		-0.721	1	
Model 4						
MA Intercept	1					
MA Log Time	-0.464	1				
EF Intercept	0.711	-0.2		1		
EF Log Time	-0.43	0.281		-0.735	1	

Note. *ns* = non-significant. Remaining correlations significant at $p < 0.05$.

Table 15

Correlations between Math Achievement and Working Memory School-Level Random Effects - Adjusted Models

	MA Intercept	MA Log Time	MA Drop	EF Intercept	EF Log Time	EF Drop
Model 1						
MA Intercept	1					
MA Log Time	-0.559	1				
MA Drop	0.29	-0.931	1			
EF Intercept	0.622	-0.178	0.029*ns*	1		
EF Log Time	-0.202	0.331	-0.296	-0.527	1	
EF Drop	-0.023*ns*	-0.296	0.33	0.19	-0.905	1
Model 2						
MA Intercept	1					
MA Log Time	-0.815	1				
EF Intercept	0.688	-0.42		1		
EF Log Time	-0.527	0.417		-0.821	1	
Model 3						
MA Intercept	1					
MA Log Time	-0.788	1				
EF Intercept	0.69	-0.395		1		
EF Log Time	-0.509	0.406		-0.798	1	
Model 4						
MA Intercept	1					
EF Intercept	0.708			1		

Note. *ns* = non-significant. Remaining correlations significant at $p < 0.05$.

Table 16 presents the results from the model of cognitive flexibility. The fixed effects show a significant main effect of log time ($b = 0.761$, $p < 0.001$), corresponding to the trajectory for a white male student with a value on the SES scale equal to zero. Blacks tend to have lower starting scores relative to whites ($b = -0.439$, $p < 0.001$), as do Hispanics ($b = -0.166$). Females tend to have higher values at the start of the series ($b = 0.174$, $p < 0.001$), as do those with higher SES ($b = 0.221$, $p < 0.001$). Trajectories for blacks tend to be more positive relative to whites ($b = 0.219$, $p < 0.001$), as is the case for Hispanics ($b = 0.121$, $p < 0.001$) and Asians ($b = 0.121$, $p < 0.001$). Females tend to have less positive trajectories ($b = -0.109$, $p < 0.001$) as do those with higher SES ($b = -0.111$, $p < 0.001$).

Based on the random effects estimates, students who have higher starting values tend to have less positive trajectories ($r = -0.868$). This is also the case for schools ($r = -0.967$). Model 2 constrains the school-level trajectory to zero given that the correlation between random effects was so close to unity. This does not impact the fixed effects results. Figure 9 displays the trajectories faceted by gender, race, and SES.

Table 17 displays the amount of reduction in variance components after including the demographics fixed effects. The student-level intercept is reduced by 4.25% while the school-level intercept is reduced by 52.83%. The trajectory variance at the student level is reduced by 3.83% while the school-level trajectory variance is reduced by 75%.

Table 16

Mixed Models of Cognitive Flexibility - with Demographics

	Model 1		Model 2	
Effect	Estimate	SE	Estimate	SE
Fixed Effects				
Intercept	-0.518***	0.0177	-0.516***	0.0167
Time (Log)	0.761***	0.0127	0.76***	0.0125
Black	-0.439***	0.0314	-0.447***	0.0302
Hispanic	-0.166***	0.0256	-0.17***	0.0248
Asian	-0.036	0.0382	-0.04	0.0373
Other Race	0.013	0.041	0.017	0.0407
Female	0.174***	0.0179	0.176***	0.0179
SES	0.221***	0.0126	0.227***	0.0123
Time (Log) X Black	0.219***	0.0236	0.228***	0.0233
Time (Log) X Hispanic	0.121***	0.0195	0.123***	0.0193
Time (Log) X Asian	0.121***	0.0312	0.122***	0.0311
Time (Log) X Other Race	0.006	0.0331	0.001	0.0331
Time (Log) X Female	-0.109***	0.0146	-0.11***	0.0146
Time (Log) X SES	-0.111***	0.0098	-0.117***	0.0098
Random Effects				
Student - Var(Intercept)	0.789		0.801	
Student - Cor(Intercept, Log Time)	-0.868		-0.872	
Student - Var(Log Time)	0.301		0.304	
School - Var(Intercept)	0.05		0.027	
School - Cor(Intercept, Log Time)	-0.967		--	
School - Var(Log Time)	0.004		--	
Residual	0.371		0.371	
Percentage Variance at School Level				
Intercept	6.0%		3.3%	
Log Time	1.3%		--	
Model Fit				
-2 Res Log Likelihood	132931		132965	
Δ -2 Res Log Likelihood	--		34	
AIC (Smaller is Better)	132945		132975	
Δ AIC	--		30	
BIC (Smaller is Better)	132978		132998	
Δ BIC	--		20	

Note. *** $p < 0.001$.

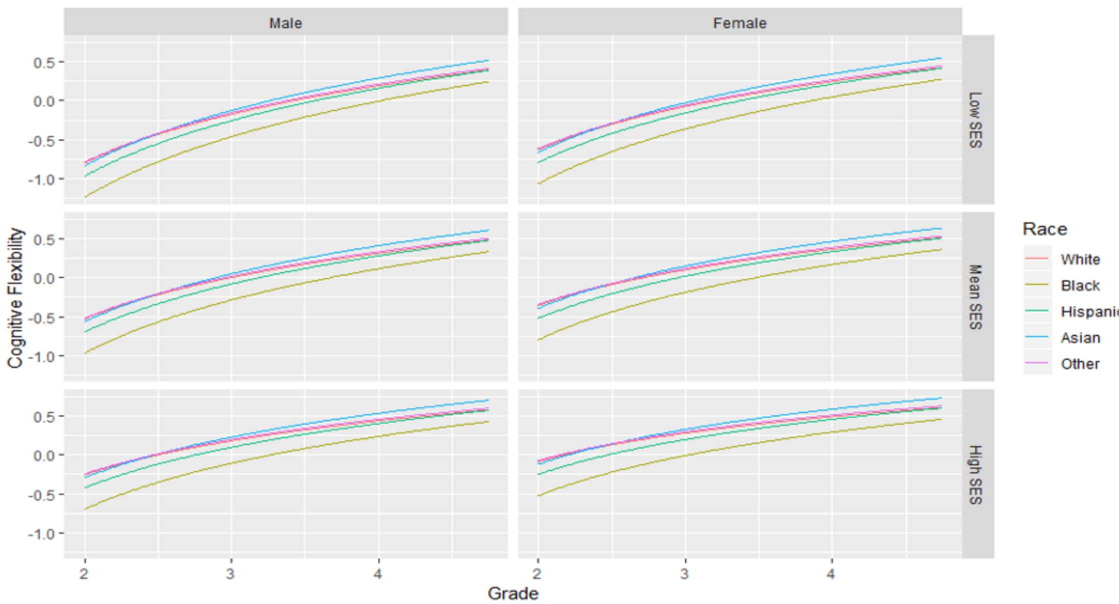

Figure 9: Model trajectories of cognitive flexibility by race, gender, and SES. Low SES defined as 1.5 standard deviations below the mean. High SES defined as 1.5 standard deviations above the mean. In the statistical model, time was rescaled so that grade 2 =zero.

Table 17
Variance Explained - Cognitive Flexibility

Model	No Demographics	With Demographics	% Explained
Model 1			
Student Level			
Student - Var(Intercept)	0.824	0.789	4.248
Student - Var(Log Time)	0.313	0.301	3.834
School Level			
School - Var(Intercept)	0.106	0.05	52.830
School - Var(Log Time)	0.016	0.004	75.000
Model 2			
Student Level			
Student - Var(Intercept)	0.856	0.801	6.425
Student - Var(Log Time)	0.328	0.304	7.317
School Level			
School - Var(Intercept)	0.044	0.027	38.636

Table 18 presents the correlations between cognitive flexibility and math achievement random effects at the student level. Higher cognitive flexibility intercepts are negatively related to math achievement trajectories (*r* = -0.265). Both intercepts are strongly, positively related (*r* = 0.539).The cognitive flexibility trajectory is negatively associated with math achievement starting values (*r* = -0.451) and positively related to math achievement trajectories (*r* = 0.235).

Table 18

Correlations between Math Achievement and Cognitive Flexibility Student-Level Random Effects - Adjusted Models

Random Effect	MA Intercept	MA Log Time	EF Intercept	EF Log Time
Model 1				
MA Intercept	1			
MA Log Time	-0.548	1		
EF Intercept	0.539	-0.265	1	
EF Log Time	-0.451	0.235	-0.936	1
Model 2				
MA Intercept	1			
MA Log Time	-0.558	1		
EF Intercept	0.537	-0.259	1	
EF Log Time	-0.453	0.232	-0.941	1

Note. All correlations significant at $p < 0.05$.

Table 19 turns to correlations in the school level random effects. The cognitive flexibility intercept is negatively associated with the math achievement trajectory (*r* = -0.211). Both intercepts are positively related (*r* = 0.525). The cognitive flexibility trajectories tend to be weaker when math achievement scores start higher (*r* = -0.527). The two trajectories are positively associated at the school level (*r* = 0.213).

Table 19

Correlations between Math Achievement and Cognitive Flexibility School-Level Random Effects - Adjusted Models

Random Effect	MA Intercept	MA Log Time	EF Intercept	EF Log Time
Model 1				
MA Intercept	1			
MA Log Time	-0.632	1		
EF Intercept	0.525	-0.211	1	
EF Log Time	-0.527	0.213	-0.999	1
Model 2				
MA Intercept	1			
EF Intercept	0.516		1	

Note. All correlations significant at $p < 0.05$.

To conclude this chapter, one final analysis is presented. The prior models for this research question estimated effects on math achievement and EF separately in order to correlate the random effects between outcomes. However, this does not answer how much EF affects math achievement directly. To answer this question, Tables 20 and 21 present results for models of math achievement that include, in addition to demographics, the EF measures of working memory and cognitive flexibility, respectively. The fixed effects estimates of EF represent how much a unit increase in EF affects math achievement. The model fit section of the tables shows the change in the model fit statistics relative to a model without the EF measure. (For DCCS, this is the math achievement model fit only to data beginning in grade 2).

The model that includes working memory as a fixed effect is the following:

$$MA_{tis} = b_0 + b_1 ln(time)_{tis} + b_2 drop_{tis} + b_3 gender_i + b_4 black_i + b_5 hispanic_i + b_6 asian_i + b_7 other\ race_i + b_8 ses_i + b_9 gender_i ln(time)_{tis} + b_{10} black_i ln(time)_{tis} + b_{11} hispanic_i ln(time)_{tis} + b_{12} asian_i ln(time)_{tis} + b_{13} other\ race_i ln(time)_{tis} + b_{14} ses_i ln(time)_{tis} + b_{15} gender_i drop_{tis} + b_{16} black_i drop_{tis} + b_{17} hispanic_i drop_{tis} + b_{18} asian_i drop_{tis} + b_{19} other\ race_i drop_{tis} + b_{20} ses_i drop_{tis} + b_{21} working\ memory_{tis} + u_{0s} + u_{0is} + u_{1s} ln(time)_{tis} + u_{1is} ln(time)_{tis} + u_{2s} drop_{tis} + u_{2is} drop_{tis} + r_{tis}$$

For DCCS, the model will be the following:

$$MA_{tis} = b_0 + b_1 ln(time)_{tis} + b_2 drop_{tis} + b_3 gender_i + b_4 black_i + b_5 hispanic_i + b_6 asian_i + b_7 other\ race_i + b_8 ses_i + b_9 gender_i ln(time)_{tis} + b_{10} black_i ln(time)_{tis} + b_{11} hispanic_i ln(time)_{tis} + b_{12} asian_i ln(time)_{tis} + b_{13} other\ race_i ln(time)_{tis} + b_{14} ses_i ln(time)_{tis} + b_{15} DCCS_{tis} + u_{0s} + u_{0is} + u_{1s} ln(time)_{tis} + u_{1is} ln(time)_{tis} + r_{tis}$$

Table 20 shows a significant effect of working memory across model specifications. A one unit increase in working memory is associated with a 0.07 unit increase in math achievement ($p < 0.001$). This results in a substantial improvement to model fit. As Model 1 shows, both BIC and AIC – presented in smaller-is-better form – decrease by 2,396.

Table 20

Mixed Models of Math Achievement Including Working Memory

Effect	Model 1 Estimate	SE	Model 2 Estimate	SE	Model 3 Estimate	SE	Model 4 Estimate	SE
Fixed Effects								
Intercept	-1.5***	0.0097	-1.498***	0.0096	-1.496***	0.0095	-1.497***	0.0087
Time (Log)	1.684***	0.0117	1.681***	0.0097	1.677***	0.0087	1.677***	0.0083
Summer Drop	-0.041***	0.004	-0.041***	0.0032	-0.04***	0.0029	-0.039***	0.0029
Black	-0.141***	0.0164	-0.142***	0.0162	-0.142***	0.016	-0.118***	0.0153
Hispanic	-0.199***	0.0132	-0.198***	0.0131	-0.198***	0.0129	-0.197***	0.0125
Asian	0.012	0.0193	0.01	0.0191	0.011	0.0188	0.003	0.0183
Other Race	-0.026	0.0202	-0.022	0.0202	-0.023	0.0199	-0.026	0.0199
Female	0.019*	0.0087	0.018*	0.0088	0.018*	0.0086	0.018*	0.0087
SES	0.224***	0.0064	0.228***	0.0063	0.228***	0.0063	0.24***	0.0061
Time (Log) X Black	-0.137***	0.0195	-0.13***	0.0174	-0.129***	0.0155	-0.147***	0.0151
Time (Log) X Hispanic	0.085***	0.0157	0.083***	0.0144	0.085***	0.0128	0.085***	0.0125
Time (Log) X Asian	-0.006	0.0251	0.001	0.0234	-0.003	0.0214	0.003	0.0211
Time (Log) X Other Race	0.001	0.0242	-0.013	0.0242	-0.012	0.0215	-0.009	0.0215
Time (Log) X Female	-0.054***	0.0104	-0.052***	0.0106	-0.052***	0.0094	-0.051***	0.0095
Time (Log) X SES	-0.064***	0.0076	-0.078***	0.0072	-0.078***	0.0065	-0.089***	0.0064
Summer Drop X Black	0.011	0.0067	0.008	0.0059	0.008	0.0054	0.008	0.0054
Summer Drop X Hispanic	-0.024***	0.0054	-0.023***	0.0049	-0.024***	0.0045	-0.024***	0.0045
Summer Drop X Asian	0.024**	0.0089	0.022**	0.0082	0.023**	0.0078	0.023**	0.0077
Summer Drop X Other Race	0.001	0.0084	0.006	0.0084	0.006	0.0078	0.006	0.0078
Summer Drop X Female	-0.004	0.0036	-0.005	0.0037	-0.005	0.0034	-0.005	0.0034
Summer Drop X SES	0.018***	0.0027	0.024***	0.0025	0.024***	0.0023	0.024***	0.0023
Working Memory	0.066***	0.0013	0.067***	0.0013	0.069***	0.0013	0.069***	0.0013

Table 20 (*continued*)

Random Effects				
Student - Var(Intercept)	0.204	0.206	0.196	0.204
Student - Cor(Intercept, Log Time)	-0.404	-0.404	-0.451	-0.484
Student - Cor(Intercept, Drop)	0.173	0.199		
Student - Var(Log Time)	0.098	0.126	0.035	0.041
Student - Cor(Log Time, Drop)	-0.863	-0.878		
Student - Var(Drop)	0.004	0.008		
School - Var(Intercept)	0.024	0.021	0.021	0.011
School - Cor(Intercept, Log Time)	-0.556	-0.737	-0.718	
School - Cor(Intercept, Drop)	0.313			
School - Var(Log Time)	0.036	0.005	0.005	
School - Cor(Log Time, Drop)	-0.926			
School - Var(Drop)	0.004			
Residual	0.048	0.048	0.05	0.05
Percentage Variance Explained at School Level				
Intercept	10.50%	9.30%	9.70%	5.10%
Log Time	26.90%	3.80%	12.50%	--
Summer Drop	50%	--	--	--
Model Fit				
-2 Res Log Likelihood	42994	43464	44082	44545
Δ -2 Res Log Likelihood from prior model		470	618	463
Δ -2 Res Log Likelihood model without WM	-2396	-2459	-2541	-2553
AIC (Smaller is Better)	43020	43484	44096	44555
Δ AIC from prior model		464	612	459
Δ AIC from model without WM	-2396	-2459	-2541	-2553
BIC (Smaller is Better)	43081	43531	44129	44579
Δ BIC from prior model		450	598	450
Δ BIC from model without WM	-2396	-2459	-2540	-2553

Note. *** $p < 0.001$. ** $p < 0.01$. * $p < 0.05$.

Table 21 presents the results from fitting the model using cognitive flexibility as the measure of EF and compares the model fit to what was obtained from a math achievement model fit without DCCS to the same data. The results are similar to what was observed for working memory. A unit increase in DCCS is associated with a significant increase in math achievement ($b = 0.032$, $p < 0.001$). The improvement in model fit is also large. Both the AIC and BIC decreased by 478 relative to the model without cognitive flexibility. Note that the magnitude of this change is smaller vis-à-vis working memory due to the smaller number of data points.

Table 21

Mixed Models of Math Achievement Including Cognitive Flexibility

Effect	Model 1 Estimate	SE	Model 2 Estimate	SE
Fixed Effects				
Intercept	0.228***	0.0087	0.227***	0.0083
Time (Log)	0.825***	0.0046	0.825***	0.004
Black	-0.274***	0.0154	-0.267***	0.0149
Hispanic	-0.163***	0.0124	-0.166***	0.0121
Asian	0.06***	0.0178	0.066***	0.0174
Other Race	-0.004	0.0193	-0.009	0.0193
Female	-0.041***	0.0084	-0.04***	0.0085
SES	0.192***	0.006	0.195***	0.0059
Time (Log) X Black	-0.048***	0.008	-0.054***	0.0072
Time (Log) X Hispanic	0.022***	0.0065	0.025***	0.006
Time (Log) X Asian	0.032**	0.0102	0.026**	0.0096
Time (Log) X Other Race	-0.018	0.0103	-0.013	0.0102
Time (Log) X Female	-0.038***	0.0045	-0.039***	0.0045
Time (Log) X SES	-0.003	0.0032	-0.006*	0.003
DCCS	0.032***	0.0014	0.032***	0.0014
Random Effects				
Student - Var(Intercept)	0.215		0.218	
Student - Cor(Intercept, Log Time)	-0.476		-0.49	
Student - Var(Log Time)	0.029		0.032	
School - Var(Intercept)	0.014		0.01	
School - Cor(Intercept, Log Time)	-0.586			
School - Var(Log Time)	0.003			
Residual	0.033		0.033	
Percentage Variance Explained at School Level				
Intercept	6.1 %		4.4 %	
Log Time	9.4 %		--	
Model Fit				
-2 Res Log Likelihood	20765		20943	
Δ -2 Res Log Likelihood from prior model			178	
Δ -2 Res Log Likelihood model without WM	-478		-480	
AIC (Smaller is Better)	20779		20953	
Δ AIC from prior model			174	
Δ AIC from model without DCCS†	-478		-480	
BIC (Smaller is Better)	20812		20976	
Δ BIC from prior model			164	
Δ BIC from model without DCCS†	-478		-480	

Note. *** $p < 0.001$. ** $p < 0.01$. * $p < 0.05$. † MA model fit to DCCS subsample.

Research Question 5: *To what degree does math achievement growth vary among schools, controlling for demographics?*

Table 5 above summarized the change in variance components from the model without demographics to the model that includes them. Student characteristics – gender, race, and SES – account for quite a bit of the between-school variability in math achievement both in terms of starting levels (66.23% of variance explained) as well as for log growth (17.02% of variance explained). To visualize this change, Figure 10 presents boxplots of the predicted school-level random effects. Focusing on Model 1, the variability in the boxplots clearly decreases for the predicted random intercepts. The range of predicted values is also reduced for log time, with fewer predictions at the very high end of the distribution. The upshot is that school quality continues to matter, especially for trajectories over time. However, a sizable proportion of differences in school outcomes can be attributed to the demographic profile of the students attending.

The correlations between math achievement and executive functioning that were reported above lead to the expectation that similar patterns will emerge for the EF measures. Indeed, as was shown above in Table 13, demographics explained a substantial percentage of between-school differences in student scores in kindergarten (75.2%) and a smaller but still notable percentage for variability in trajectories (26.47%). Figure 11 visualizes these changes. Looking at Model 1, the spread in the intercept random effects becomes much smaller after adjusting for controls. For log trajectories, the width of the box becomes smaller, meaning that the interquartile range is more constrained around the median. School quality matters for executive functioning over time, but much of the between-school differences are again attributable to the student milieu in those schools.

Table 17 above shows the amount of variance explained for the cognitive flexibility measure, and Figure 12 provides boxplot summaries. Demographics again are closely related to school-level variance. Including the student-level characteristics in the model reduces the variance component for the intercept by over half (52.83%) and the variance component for the log trajectory by three-quarters (75%). These results are based on a much shorter timeframe vis-à-vis the math achievement and working memory measures, but the story is similar to what was seen for math achievement and working memory.

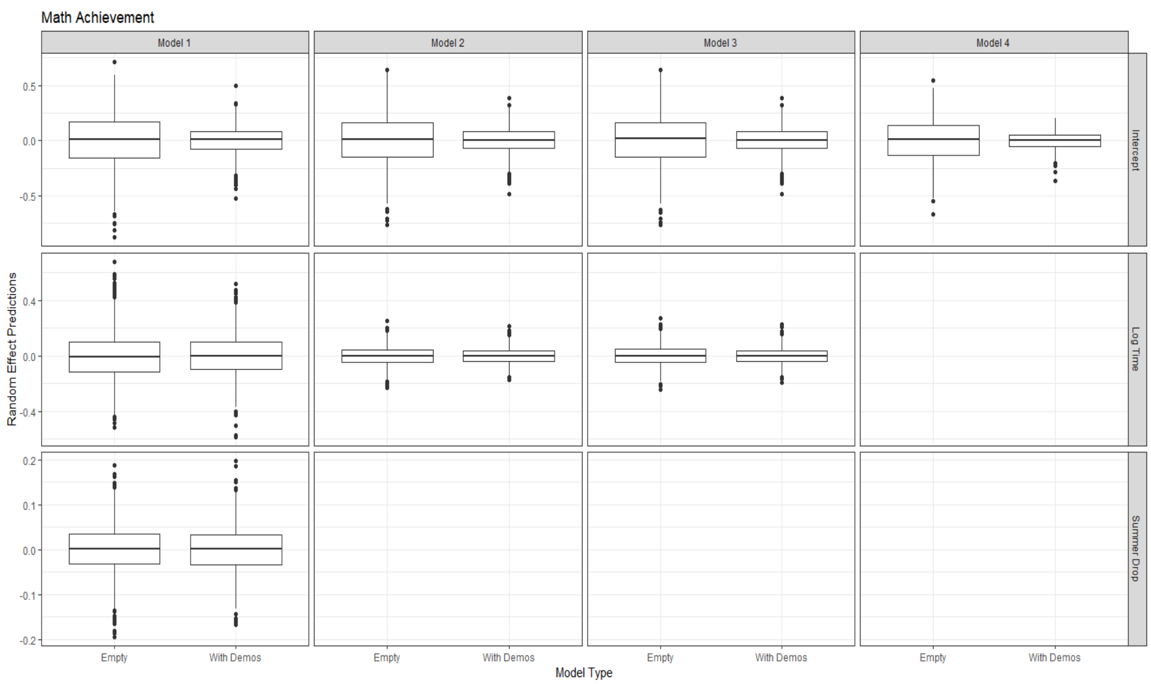

Figure 10: Boxplots of school-level random effects predictions (BLUPs) for math achievement by model. Boxplot on left of each panel shows distribution of random effects from model that does not adjust for demographics, boxplot on right shows distribution after controlling for demographics.

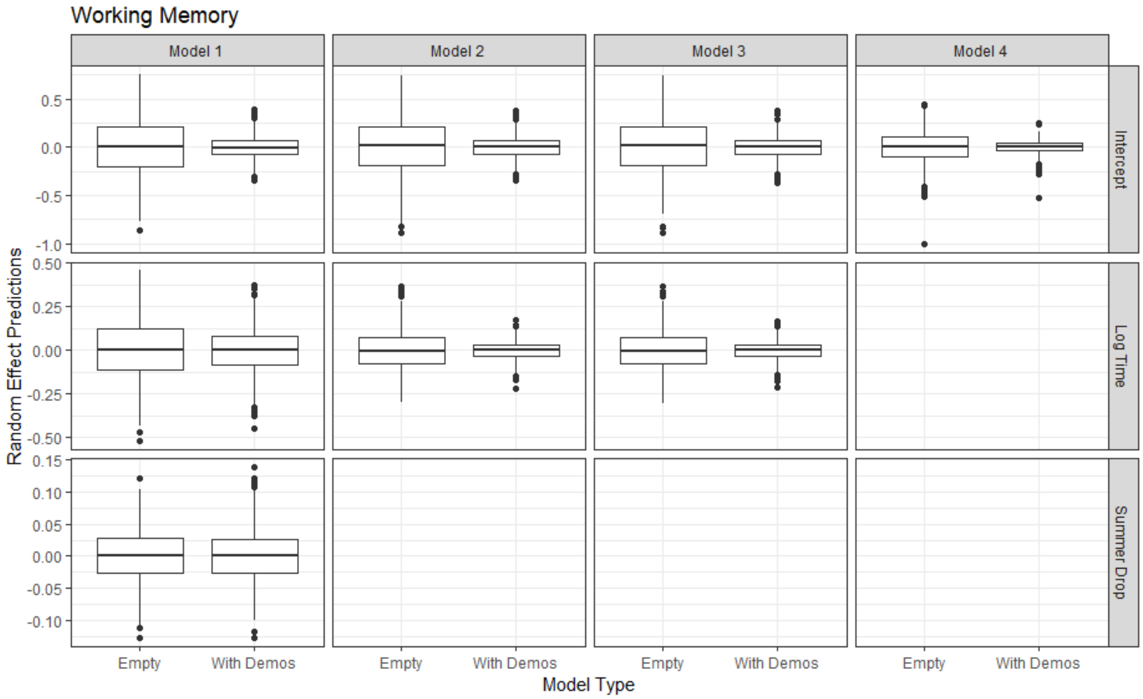

Figure 11: Boxplots of school-level random effects predictions (BLUPs) for working memory by model. Boxplot on left of each panel shows distribution of random effects from model that does not adjust for demographics, boxplot on right shows distribution after controlling for demographics.

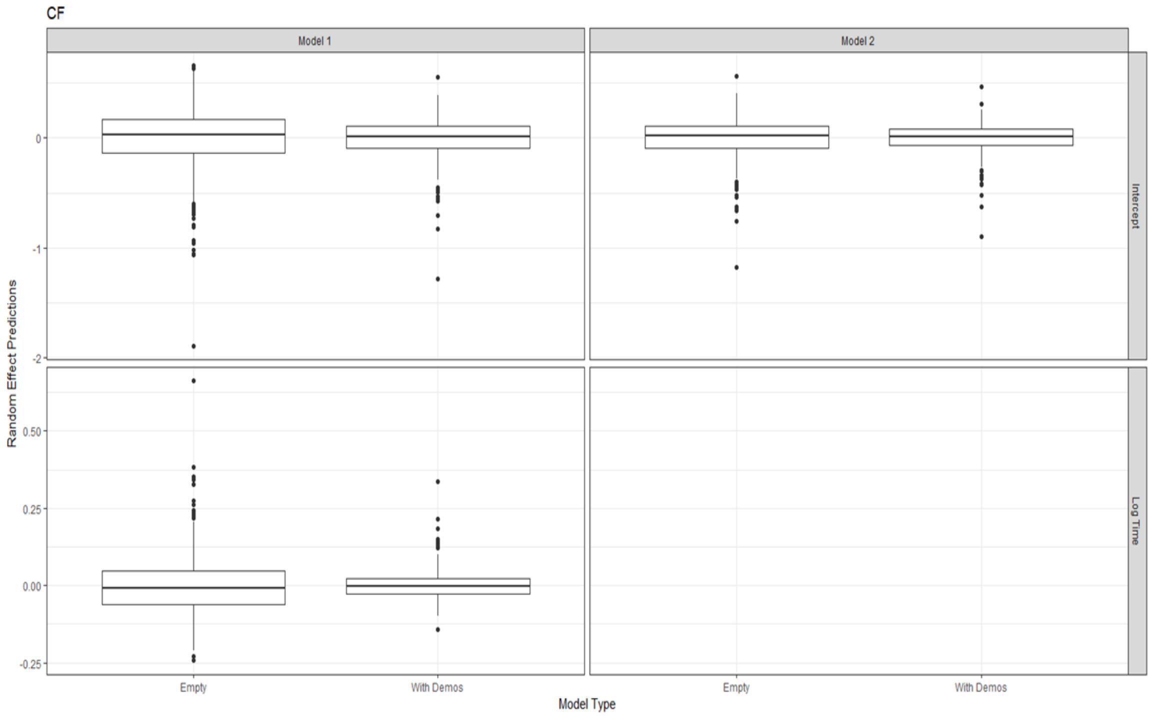

Figure 12: Boxplots of school-level random effects predictions (BLUPs) for cognitive flexibility by model. Boxplot on left of each panel shows distribution of random effects from model that does not adjust for demographics, boxplot on right shows distribution after controlling for demographics.

CHAPTER 5

DISCUSSION

The results of the analysis conducted for this study are discussed in this chapter, and conclusions are drawn based on these findings. First is a summary, along with a discussion on how the results answer each of this study's four research questions and how these results support or fail to support the findings of previous work conducted in this field. This is followed by a discussion of the limitations present within the study, along with possibilities for future research. Finally, a series of conclusions are delineated based on the findings.

Summary of Results

This study focused on the examination of the following research questions:

Research Question 1: *What is the average math achievement growth trajectory from kindergarten through fourth grade, accounting for the nesting of students within schools, and summer loss?*

The analyses conducted in relation to this research question found a positive and significant trend over time after accounting for a significant decrease at the start of the fall semester and school and student random effects. Additionally, students with lower scores in kindergarten generally had trajectories that were more positive, while those with below average starting points had a more negative drop in summer. This drop was more pronounced among those with above average trajectories.

At the school level, lower average starting values were associated with above average trajectories and larger decreases in summer, and above average trajectories were associated with more negative drops in summer. These results departed from that of previous literature, which found that early mathematics difficulties were associated with the lowest growth rates in later grades (Morgan, Farkas, & Wu, 2009), thought it supported the lag model as opposed to the cumulative growth model (Morgan, Farkas & Wu, 2009). The shape of the math achievement trajectory was nonlinear, with the strongest average improvements occurring in early grades and weaker improvements in each successive year.

Research Question 2: *To what degree do gender, racial, or socioeconomic status differences account for variation in the intercepts, slopes, and summer drops in the mathematics trajectories at the student and school levels?*

With respect to the second research question, these models also found a significant increase in math achievement over time and a significant decrease over summer. The fixed effects results found significantly reduced scores for Blacks and Hispanics and significantly increased scores for females and those with higher SES at the outset of the series. The increase in scores over time was reduced for Blacks, females, and those with higher SES, and slightly increased for Hispanics, while the summer drop was more pronounced for Hispanics and less pronounced among Blacks, Asians, and those with higher SES.

Based on the random effects, higher student starting scores were associated with reduced trajectories and less pronounced drops during summer, while there was also a large, negative association between trajectory and summer drop. Additionally, schools with higher starting values had lower trajectories and less pronounced summer drops, with the association between summer drop and trajectory being strong and negative. Comparing the variance components between the models with and without demographics found that, after controlling for demographics, most of the variability in intercepts, trajectories, and summer drop was at the student level.

Research Question 3: *Is executive functioning associated with change in math achievement over this period?*

With respect to the third research question, the models of working memory and cognitive flexibility found a significant, positive trajectory with a significant summer drop. Higher starting scores were again associated with reduced trajectories and a less pronounced effect of the summer drop. A strong, negative association was found between trajectories and summer drop. Most of the variability in working memory scores in kindergarten, trajectories, and summer drop existed at the student level.

The random effects from the working memory and cognitive flexibility were then predicted and correlated with the predicted random effects from the math achievement model. This analysis found a strong, positive correlation between kindergarten working memory and math achievement, with negative, weaker associations found between kindergarten working memory and math achievement trajectories and summer drops. Higher working memory trajectories were associated with improved math trajectories and worse summer drops. Less pronounced summer drops for working memory were associated with reduced math intercepts, reduced math trajectories, and less pronounced summer drops. Similar results were found at the school level. These results answer the research question affirmatively: math achievement and working memory tend to start and evolve together.

Cognitive flexibility was found to significantly and substantially improve over time, with higher starting scores being strongly associated with reduced trajectory at both the student and school levels. When correlating the random effects from the cognitive flexibility model with the random effects predicted from the math achievement model, a strong association was found between math achievement and cognitive flexibility at baseline. A negative association was found between the cognitive flexibility intercept and math achievement trajectory. A negative and moderate association was indicated between the math achievement intercept and executive functioning trajectories, with these two trajectories themselves being positively related. Similar results were found at the school level. These results again show that executive function is related to math achievement.

Research Question 4: *Which student sociodemographic characteristics are most strongly associated with change in math achievement, and to what degree does controlling for student sociodemographic characteristics alter the association between executive functioning and change in math achievement?*

Regarding the fourth and final research question, a significant, positive effect was found over time for both working memory and cognitive flexibility, with a significant, negative summer drop, after controlling for demographics. Kindergarten working memory scores were significantly reduced for Blacks, Hispanics, and Asians, and were increased for females and those with higher SES. Trajectories were increased for Hispanics relative to whites and were reduced for those with higher SES. The summer drop was also found to be less negative among Blacks, those of other race, and those with higher SES, with the effect being more pronounced for females. Higher intercepts were also associated with lower trajectories along with less pronounced summer drops, with the association between trajectory and summer drop being negative and strong, with these patterns also appearing at the school level. Demographic fixed effects were found to reduce school-level variance more than student-level variance, as was found in the case of math achievement.

Next, at the student level, a strong, positive association was found between the working memory and math achievement intercepts after controlling for demographics. The working memory intercepts had a minimal association with math achievement trajectory along with a small, negative association with summer drop. The working memory trajectories had a weak, negative association with math achievement intercept along with math achievement summer drop. The association was positive for the math achievement trajectory. Additionally, summer drop in working memory had a negative association with the math achievement intercept and trajectory. The association with math achievement summer drop was positive. School-level results were similar to those found at the student-level. However, while the size of the correlations at the student-level were similar between the adjusted and unadjusted models, the correlations at the school-level tended to be smaller in the adjusted model. That is, although math achievement and working memory trajectories tend to move together, the size of the association for schools is diminished once one controls for demographics. This result coincides with the finding that demographics reduces more school-level variability than student-level variance.

Regarding cognitive flexibility, significant improvement was again found over time, with starting scores being reduced for Blacks and Hispanics as compared with Whites. Starting scores were also higher for females and those with higher SES. Trajectories were more positive for Blacks, Hispanics, and Asians as compared with Whites, with trajectories being reduced for females and those with higher SES. Students with higher starting values tended have lower trajectories. The inclusion of demographics fixed effects reduces the variance in intercepts at the student level minimally, while over 50% of the school level variance was accounted for once demographics were included. The student-level trajectory variance is also reduced minimally, with the school-level trajectory variance was reduced by 75%.

Next, a strong, positive association was found between the cognitive flexibility and math intercepts, with higher cognitive flexibility intercepts being negatively and weakly associated with math trajectories. Additionally, the cognitive flexibility trajectory is negatively associated with math starting values, and positively associated with math trajectories. The school-level results mirrored these findings. The inclusion of demographics did less to affect the school-level variance components than was the case with working memory. This may be due to the reduced time period available for cognitive flexibility, as improvements in math achievement and executive functioning (measured with working memory) both showed the greatest acceleration in prior grades.

Research Question 5. *To what degree does math achievement growth vary among schools, controlling for demographics?*

Examining the amount of school-level variability in math achievement explained by demographics revealed that student-level characteristics (gender, race, SES) accounted for a great deal of between-school differences. Indeed, the variance components at the school level were reduced much more than the student-level variance components. This is indicative of the importance of the school's overall demographic make-up for math achievement outcomes, both in kindergarten and the over-time trajectories. A similar story was found for both executive functioning outcomes. Nonetheless, school demographics did not account for all between-school differences, which leaves a role for school quality above and beyond the make-up of its students.

Discussion

Regarding the answers to these research questions, first, with respect to Research Question 1, excluding demographics, a significant growth trajectory was found ($b = 1.751$, $p < 0.001$) after accounting for summer drop and school and student random effects. With demographics, this effect is nearly the same. In addition, results found this trajectory to not be linear but to decrease over time, with larger and larger decreases evident year after year.

In relation to this study's second research question, intercepts were decreased for Blacks and Hispanics, but were increased for females and those with higher SES. The results for race corresponded with previous literature, which found minorities like Hispanics and Blacks to have reduced mathematics performance as compared with Whites and Asians (Sonnenschien & Sun, 2017). This may be explained by Blacks and Hispanics having reduced working memory and cognitive flexibility as compared with Whites (Little, 2017), and may also suggest the continuing importance of the parents in explaining racial differences in mathematics skills (Sonnenschien & Sun, 2017).

The SES result also corresponded with past literature, with children coming from low-income households having poorer math skills compared to those from high-income households (Morgan, Farkas, and Wu, 2009). This may be explained by those with higher SES having improved working memory and cognitive functioning (Little, 2017) and as those from lower SES backgrounds have fewer opportunities to learn (Morgan, Farkas, & Wu, 2009) and also generally have reduced EF (Zelazo, 2015).

The improved scores for females are supported by some previous literature (Halpern et al., 2007), but not others (Reilly, Neumann, & Andrews, 2015). These results for gender do not indicate any support for the perceived disadvantages among females leading to poorer mathematics achievement and speak to the apparent complexity pertaining to the association between gender and achievement. They do, however, fail to support earlier research which has suggested that anxiety among females may lead to reduced mathematics performance (Lent et al., 1991; Passolunghi et al., 2016; Peters, 2015), and also fail to support the explanation of stereotype threat as the basis behind gender differences in mathematics performance (Ganley et al., 2013).

Additionally, the mathematics performance trajectory was less positive for Blacks, slightly more positive for Hispanics, slightly less positive for females, and less positive among those with higher SES. The higher scores among females at the outset combined with a reduced trajectory departs from previous literature which found gender differences to significantly increase over time (Levine et al., 1999; Reilly, Neumann, & Andrews, 2015). The effect of summer drop was more negative among Hispanics, but less negative among Blacks, Asians, and those with higher SES. More closely answering this research question with respect to the reduction in variance, the addition of these demographic measures had a substantially greater impact in reducing school-level variances as compared with student-level variances. Specifically, the introduction of these demographics reduced the variance of the student-level intercept by 8.11%, while variance in the school-level intercept was reduced by 66.23%. With respect to slopes, student-level variance was reduced by 4.55%, while school-level variance was reduced by 17.02%. Regarding variance in the summer drop, this was unchanged at the student level, but reduced by 20% at the school level.

Regarding the third research question, scores in kindergarten on working memory had a strong association with scores in kindergarten on math achievement ($r = 0.711$). Higher initial scores on working memory were also associated with weaker math trajectories ($r = -0.105$) along with more severe summer drops ($r = -0.081$). While improved working memory trajectories were associated with improved math achievement trajectories ($r = 0.238$), they were also associated with more severe summer drops ($r = -0.200$). Less severe working memory summer drops were also associated with reduced math achievement intercepts ($r = -0.25$), reduced math trajectories ($r = -0.185$), and less severe math summer drops ($r = 0.257$). Results for school-level random effects were found to be very similar.

A strong correlation was found between math scores and cognitive flexibility at the student level ($r = 0.522$), with a negative, significant association found between the cognitive flexibility intercept and the math trajectory ($r = -0.255$). A moderate, negative association was found between the math intercept and executive function trajectory ($r = -0.469$), with both trajectories positively but weakly associated ($r = 0.230$). Again, school-level random effects results were very similar.

These results concorded with that of previous literature. First, previous research has found a positive relationship between EF and academic achievement (Monette, Bigras & Guay, 2011; Morgan, et al., 2017), with EF being composed of the skills of working memory, cognitive flexibility, and inhibitory control (Monette, Bigras & Guay, 2011; Zelazo, 2015). Other research has found EF skills to positively impact goal attainment, which contributes to academic success (Little, 2017), while others have found a direct, positive link between EF and mathematics performance (Blair and Razza, 2007; Bull, Espy & Wiebe, 2008; Bull & Scerif, 2011; Clark, Pritchard & Woodward, 2010; Monette, Bigras & Guay, 2011; Morgan et al., 2016), especially working memory (Monette, Bigras & Guay, 2011), with deficits in EF associated with deficits in mathematics performance (Morgan, et al., 2017). All of these previous results were upheld in this study.

With regard to the fourth research question, trajectories were found to be more positive for Hispanics as compared with Whites, with higher SES being associated with lower trajectories. The summer drop was less severe for Blacks, those of other race, and those with higher SES, with a more severe drop for females.

With respect to Research Question 5, sociodemographics accounted for a large degree of between-school differences, with the variance components primarily being reduced at the school level. This suggests the importance of the school's demographic composition in math achievement outcomes, with the results suggesting similar importance in relation to executive functioning outcomes. With respect to executive functioning, Morgan et al. (2017, 2019) found that reduced executive functioning predicts mathematics difficulties later in schooling. This finding was confirmed in the present study, while this study also explored how sociodemographics at the individual as well as the school level serve to impact math trajectories and the role of EF in mathematics achievement over time.

Limitations

Several limitations are present in this study. First, limitations are present based on the method of sampling that was used in the ECLS-K methodology. These data can be used to produce estimates that are nationally representative of school and teacher characteristics, but only regarding the kindergarten, or base-year, data. These data do not allow for nationally representative estimates to be calculated in grades later than kindergarten. This limits the generalizability and external validity of the results obtained. Specifically, the results relating to the models included within this study did not only incorporate data from the kindergarten year but include the following years of data for the purposes of modeling trajectories as well as changes in the trajectory and impacts that other variables have on the trajectories. Based on the inclusion of this larger set of data, the results cannot be generalized to the larger population except for the results obtained specifically on the initial kindergarten data. Due to this, any other results obtained can only tentatively be generalized to the larger population, which serves to limit the external validity of the study and the results obtained.

A second limitation of this study is that since it is based on observational data, it is very difficult to test for causality. Because of this, whether any variable has a causal impact on any other variable cannot be definitively determined based on the results obtained in this study. For this to be determined, an alternate methodology would need to be used, such as a cross-lagged panel structural equation modeling. However, these are difficult to implement with random effects, while the random effects were essential to addressing the research questions.

Third, despite being nationally representative of at least the kindergarten level, the results found cannot be applied to any country other than the United States or any time period other than that included within the data used for this study. While these results pertaining to the kindergarten data specifically can be generalized to that of the larger population, the results cannot be expected to hold true for any other country or any other time period before or after that which was included within the study's data. The extent to which these relationships hold or do not hold in these other locations or time periods cannot be determined based on this study’s results.

Fourth, this study was also limited because of the number and nature of the variables included in the analyses conducted. Regarding outcomes, math achievement was the primary focus. Therefore, these results do not speak to the impact of these same independent variables on other areas of achievement scholastically, or about other important outcomes such as reading. Similarly, with respect to the independent variables included within this study, this set of predictors were also fairly limited in scope. The variables that were included in the model were chosen to be consistent with prior theory. However, the extent to which other important demographic variables, other variables associated with executive functioning, or measures relating to other forms of cognitive processing may impact this or other outcomes is unknown.

Future Research

These limitations can also be drawn upon in order to provide recommendations for future research. First, the sampling methodology associated with the data analyzed in this study was described as a limitation, with only the kindergarten-level data being nationally representative and allowing for only these results obtained to be generalized to the larger population. Since these data were only nationally representative at the kindergarten level, as multiple years of data were incorporated into these analyses, this indicates that the results obtained in the study in the models conducted cannot be generalized to any larger population, which limits the external validity associated with this study. Future research using data which is nationally representative in all years of administration would allow for the generalization of all results to the larger population and would substantially improve the external validity of the study.

Secondly, the issue of causality was raised as another limitation of the study as the models used in this study did not allow for the determination of causality between the independent and dependent variables. Specifically, other designs may be more appropriate for testing for causality, with these including such methods as panel regression cross-lagged panel structural equation modeling, or marginal structural models. Future studies could seek to also determine causality by seeing how well the results replicate using one of these alternative designs. This would allow for expansion of this area of study and further growth through the further exploration of the relationships between these independent and dependent variables.

The third limitation discussed consisted of these data being specific to the United States and to a certain time period, and not being generalizable to any other country or time period before or after that which was included in the data analyzed in this study. Future research could examine these same relationships and use the same models while focusing upon data that was collected in another country or another time period. This would allow for a determination of whether the results found in this current study hold in other countries and cultures, while the analysis of data from other time periods would allow for the determination of whether these results hold in time periods before or after that which was focused upon in this study. This would allow for a conclusion of whether these results are generally consistent or generally inconsistent across countries, cultures and time periods.

Finally, the fourth limitation discussed related to the number and nature of the variables included in the study. The specific limitations mentioned consisted of the inclusion of only mathematics achievement as an outcome variable of interest as well as the limited scope of the independent variables included in the study. Future research could expand upon this study by examining other outcome variables above and beyond math achievements, such as reading or scientific reasoning. With respect to the independent variables, the set of demographic variables focused upon in this study could be expanded upon in future research, with additional relevant, non-demographic predictors added to the models. This would also allow for the expansion of this area of study by determining whether these same predictors impact other outcomes in the same way, as well as through the determination of how other independent variables impact mathematics achievement as well as other related outcomes.

Conclusion

The goal of this study was to examine the relationship between executive functioning and mathematics performance, expanding upon this area of literature, as well as to examine the relationship between demographic variables and mathematics performance. Findings have varied across studies, and by using a large national dataset, this study hoped to clarify some of these discrepancies. Additionally, this study aimed to add to this area of research by modeling the relationships between these variables in more complex, and more realistic ways than have been done in the past. While previous research has examined the relationship between demographics, executive functioning, and mathematics performance, these measures have rarely all been incorporated into a single cohesive study. Additionally, this study also adds to this body of literature by examining five-year growth trajectories in mathematics achievement and executive functioning together, modeling the association between the intercepts and slopes of both trajectories.

The relationship between the demographic variables of interest, executive functioning, and mathematics performance were examined comprehensively, serving to expand upon this area of research and allowing for future examination of the relationships between these measures in more complex and realistic ways. These findings also helped to clarify the relationships between these measures in cases where previous research found differing results. As the present study incorporated a large, national data set, these discrepancies in previous literature were likely due to the use of smaller, more specific data sets or due to the use of more specific models. Additionally, this study was also successful in examining the relationship between demographics, executive functioning, and mathematics performance within the context of a single study.

Regarding purely new findings, these related most strongly to the analyses conducted in relation to this study's fourth research question. Previous research was identified and discussed in this study's literature review as well as in this chapter which related to Research Questions 1 through 3. These novel findings in relation to Research Question 4 found trajectories relating to math achievement that were more positive for Hispanics as compared with Whites, and with higher SES associated with lower trajectories. These results indicate that while all races had a positive trajectory for math achievement, Hispanics were found to increase more rapidly in their achievement as compared with whites. Summer drop was found to be less severe for Blacks, as well as for those of other race, and those with higher SES, and a more severe drop found among females. These results indicate that Blacks and those of other race are impacted less negatively from the cessation of schooling over the summer, while this was also true among those of higher SES. Additionally, the cessation of schooling was also found to have a more severe impact among females as compared with males.

Implications of the Results to Educational Practice

Overall, the findings of this study highlight the importance of executive functioning as well as demographic differences in mathematics performance as well as the trajectory of mathematics performance. These findings can be implemented by school districts, schools, and teachers to create targeted programs that assist students who are more likely to have lower mathematics performance, as well as those who are more likely to improve their mathematics performance more slowly over time. Additionally, the results of the study, if summarized and made available to parents, could be used to educate them on the factors impacting mathematics performance and the improvement in mathematics performance over time. Specifically, they could be provided to them as a recommendation to spend additional time assisting their children with mathematics, or to hire a tutor, for example, if their child is among those who are predicted to have lower mathematics performance or to have reduced improvement in mathematics performance over time. Regarding the findings in relation to the summer drop, while a large body of literature already exists finding similar effects, this study helps to reinforce the importance of the negative impact of the cessation of schooling during summer. Specifically, the results of this study would strongly suggest the importance of year-round schooling, which could be achieved through the extension of the school year, or individually, by parents, through their own private and individual continuation of school through home-study, the use of tutors, etc. The importance of mathematics and a country excelling within this field was highlighted earlier in this dissertation, and all the results found can be seen as highly relevant to this issue and goal. By considering the results of this study and properly implementing programs that can help American students excel in mathematics, they will not only have brighter futures, but this will also allow America itself to continue to excel in mathematics, science, technology, and engineering on the global stage.

REFERENCES

Anderson, P. (2002). Assessment and development of executive function (EF) during childhood. *Child neuropsychology*, *8*(2), 71-82.

Ashcraft, M. H., & Krause, J. A. (2007). Working memory, math performance, and math anxiety. Psychonomic Bulletin & Review 14(2), 243-248.

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, 699-713.

Baddeley, A. (1996). Exploring the Central Executive. Journal of Experimental Psychology, 49, 5-28.

Baddeley, A. (1998). Recent developments in working memory. Current Opinion in Neurobiology, 8(2), 234-238.

Bauer, D. J., & Curran, P. J. (2005). Probing interactions in fixed and multilevel regression: Inferential and graphical techniques. Multivariate behavioral research, 40(3), 373-400.

Blair, C., & Razza, R. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. Society for Research in Child Development, 78(2), 647-663.

Brooks-Gunn, J., Klebanov, P. K., & Duncan, G. J. (1996). Ethnic differences in children's intelligence test scores: Role of economic deprivation, home environment, and maternal characteristics. Child development, 67(2), 396-408.

Bull, R., & Scerif, G. (2001). Executive functioning as a predictor of children's mathematics ability: Inhibition, shifting, and working memory. Developmental Neuropsychology, 19(3), 273-293. doi: 10.1207/s515326942DN1903_3.

Bull, R., Espy, K. & Wiebe, S. (2008). Short-term memory, working memory, and executive functioning in preschoolers: longitudinal predictors of mathematical achievement at age 7 years. Developmental Neuropsychology, 33(3), 205-228.

Carlson, S.M., Zelazo, P.D., & Faja, S. (2013). Executive function. In P.D. Zelazo (Ed.), Oxford, Handbook of developmental psychology, Vol. 1: Body and mind (pp.706-743). New York: Oxford University Press.

Clark, C., Pritchard, V. & Woodward, L. (2010). Preschool executive functioning abilities predict early mathematic achievement. Developmental Psychology, 46(5), 1176-1191.

Colom, R., Shih, P., Flores-Mendoxa, C. & Quiroga, M. (2006). The real relationship between short-term memory and working memory. Memory, 14(7), 804-813.

Colom, R., Flores-Mendoza, C., Quiroga, M. & Privado, J. (2005). Working memory and general intelligence: the role of short-term storage. Personality and Individual Differences, 39, 1006-1014.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*, 450-466.

Diamond, A. (2013). Executive functions. Annual Rev. Psychology, 64, 135-168.

Espy, K.A., McDiarmid, M.D., Cwik, M.F., Stalets, M.M., Hamby, A., & Senn, T.E. (2004). The contribution of executive functions to emergent mathematic skills in preschool children. Developmental Neuropsychology, 6, 465-486.

Fitzmaurice, G. M., Laird, N. M., & Ware, J. H. (2012). Applied longitudinal analysis (Vol. 998). John Wiley & Sons.

Ganley, C. M., Mingle, L. A., Ryan, A. M., Ryan, K., Vasilyeva, M., & Perry, M. (2013). An examination of stereotype threat effects on girls' mathematics performance. Developmental Psychology, 49(10), 1886-1897.

Ganley, C. M., Vasilyeva, M., & Dulaney, A. (2014). Spatial ability mediates the gender difference in middle school students' science performance. Child Development 85(4), 1419-1432.

Geary, D. C. (2011). Consequences, characteristics, and causes of mathematical learning disabilities and persistent low achievement in mathematics. Journal of Developmental and Behavioral Pediatrics: JDBP, 32(3), 250.

Geary, D. C. (2013). Early foundations for mathematics learning and their relations to learning disabilities. Current Directions in Psychological Science, 22(1), 23-27.

Geary, D.C. (2013). Learning Disabilities in Mathematics: Recent Advances. In H. L. Swanson, K. R. Harris, and S. Graham (Eds.), Handbook of learning disabilities (2nd ed., pp. 593-606). New York, NY: The Guilford Press. doi: 10.1016/b978-012-374748-8.00002-0

Halpern, D. F., Benbow, C. P., Geary, D. C., Gur, R. C., Hyde, J. S., & Gernsbacher, M. A. (2007). The science of sex differences in science and mathematics. Psychological Science in the Public Interest, 8(1), 1-51.

Handbook of Behavioral Neuroscience. (2016). Cognitive flexibility. Retrieved on 28 May 2018, from https://www.sciencedirect.com/topics/neuroscience/cognitive-flexibility.

Holcomb, Z. C. (2016). Fundamentals of descriptive statistics. Routledge.

Hooper, S., Roberts, J. & Sideris, J. (2010). Longitudinal predictors of reading and math trajectories through middle school for African American versus Caucasian students across two samples. Developmental Psychology, 46(5), 1018-1029.

Hyde, J. S., & Mertz, J. E. (2009). Gender, culture, and mathematics performance. Proceedings of the National Academy of Sciences, 106(22), 8801-8807. doi: 10.1073/pnas.0901265106

Jacob, R., & Parkinson, J. (2015). The potential for school-based interventions that target executive function to improve academic achievement: A review. Review of Educational Research, 85(4), 512-552.

Kirkham, N.Z., Cruess, L., & Diamond, A. (2003). Helping children apply their knowledge to their behavior on a dimension-switching task. Developmental Science, 5,449-476.

Levine, S. C., Huttenlocher, J., Taylor, A., & Langrock, A. (1999). Early sex differences in spatial skill. Developmental psychology, 35(4), 940.

Little, M. (2017). Racial and socioeconomic gaps in executive function skills in early elementary school: nationally representative evidence from the ECLS-K: 2011. Educational Researcher, 46(2), 103-109.

Maerten-Rivera, A., Myers, N., Lee, O., & Penfield, R. (2010). Student and school predictors of high-stakes assessment in science. Science Education, 94, 937–962.

Masten, A. S., Herbers, J. E., Desjardins, C. D., Cutuli, J. J., McCormick, C. M., Sapienza, J. K., Long, J.D. & Zelazo, P. D. (2012). Executive function skills and school success in young children experiencing homelessness. Educational Researcher, 41(9), 375-384.

Mather, N., &Woodcock, R.W. (2001). Examiner’s Manual: Woodcock–Johnson III Tests of Achievement. Itasca, IL: Riverside Publishing.

Mau, W.-C., & Lynn, R. (2000). Gender differences in homework and test scores in mathematics, reading and science at tenth and twelfth grade. Psychology, Evolution & Gender, 2, 119–125.

McClelland, M., Cameron, C., Duncan, R. & Bowles, R. (2014). Predictors of early growth in academic achievement: the head-toes-knees-shoulders task. Frontiers in Psychology, 5(599), 1-12.

Mezzacappa, E. (2004). Alerting, orienting, and executive attention: Developmental properties and sociodemographic correlates in an epidemiological sample of young, urban children. Child development, 75(5), 1373-1386.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. Cognitive psychology, 41(1), 49-100.

Monette, S., Bigras, M & Guay, M. (2011). The role of the executive functions in school achievement at the end of Grade 1. Journal of Experimental Child Psychology, 109, 158-173.

Morgan, P., 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.

Morgan, P. L., Li, H., Farkas, G., Cook, M., Pun, W. H., & Hillemeier, M. M. (2017). Executive functioning deficits increase kindergarten children's risk for reading and mathematics difficulties in first grade. Contemporary educational psychology, 50, 23-32.

Morgan, P. L., Farkas, G., Wang, Y., Hillemeier, M. M., Oh, Y., & Maczuga, S. (2019). Executive function deficits in kindergarten predict repeated academic difficulties across elementary school. Early Childhood Research Quarterly, 46, 20-32.

Mulligan, G. M., McCaroll, J. C., Flanagan, K. D., & Potter, D. (2018). Findings from the Fourth-Grade Round of the Early Childhood Longitudinal Study, Kindergarten Class of 2010-11 (ECLS-K: 2011). First Look. NCES 2018-094. National Center for Education Statistics.

National Mathematics Advisory Panel. (2008). Foundations for success: Final report of the National Mathematics Advisory Panel. Washington, D.C.: United States Department of Education. Retrieved from: http://www.ed.gov/about/bdscomm/list/mathpanel/report/final-report.pdf

Noble, K. G., Norman, M. F., & Farah, M. J. (2005). Neurocognitive correlates of socioeconomic status in kindergarten children. Developmental science, 8(1), 74-87.

Paglin, M., & Rufolo, A. M. (1990). Heterogeneous human capital, occupational choice, and male-female earnings differences. Journal of Labor Economics, 8(1, Part 1), 123-144.

Passolunghi, M. C., Caviola, S., De Agostini, R., Perin, C., & Mammarella, I. C. (2016). Mathematics anxiety, working memory, and mathematics performance in secondary-school children. Frontiers in psychology, 7, 42.

Peters, R. S. (2015). Ethics and Education (Routledge Revivals). Routledge.

Preacher, K. J., Curran, P. J., & Bauer, D. J. (2006). Computational tools for probing interactions in multiple linear regression, multilevel modeling, and latent curve analysis. Journal of educational and behavioral statistics, 31(4), 437-448.

Raudenbush, S. W., & Bryk, A. S. (2002). Hierarchical linear models: Applications and data analysis methods. Thousand Oaks, CA: Sage Publications.

Reilly, D., Neumann, D. L., & Andrews, G. (2015). Sex differences in mathematics and science achievement: A meta-analysis of National Assessment of Educational Progress assessments. Journal of Educational Psychology, 107(3), 645.

Rivera-Batiz, F. L. (1992). Quantitative literacy and the likelihood of employment among young adults in the United States. Journal of Human Resources, 313-328.

Robinson, J., Lubinski, S. & Copur, Y. (2011). The effects of teachers' gender stereotypical expectations on development of the math gender gap. Society for Research on Educational Effectiveness, 1-11.

Rosen, V. & Engle, R. (1997). The role of working memory capacity in retrieval. Journal of Experimental Psychology, 126(3), 211-227.

Schrank, R. (2010). Woodcock-Johnson II tests of cognitive abilities. Retrieved on 29 May 2018, from www.iapsych.com/articles/schrank2010ip.pdf

Scott, J. P. (1962). Critical periods in behavioral development. Science, 138(3544), 949-958.

Singer, F. M., Ellerton, N. F., & Cai, J. (2015). Mathematical problem posing. New York: Springer. http://doi. org/10.1007/978-1-4614-6258-3.

Sonnenschien, S. & Sun, S. (2017). Racial/ethnic differences in kindergartner's reading and math skills: parents' knowledge of children's development and home-based activities as mediators. Infant and Child Development, 26(5).

Swanson, H.L. & Alloway, T.P. (2012). Working memory, learning, and academic achievement. In Harris, K.R., S. Graham, & Urdan, T. (Eds.) APA Educational Psychology Handbook: Vol. 1. Theories, Constructs, and Critical Issues. (pp. 327-366). Washington, D.C., U.S.: American Psychological Association. doi: 10.1037/13273-012

Swanson, H., Mink, J. & Bocian, K. (1999). Cognitive processing deficits in poor readers with symptoms of reading disabilities and ADHD: more alike than different? Journal of Educational Psychology, 91(2), 321-333.

Tourangeau, K., Nord, C., Lê, T., Wallner-Allen, K., Vaden-Kiernan, N., Blaker, L. and Najarian, M. (2017). Early Childhood Longitudinal Study, Kindergarten Class of 2010–11 (ECLS-K:2011) User's Manual for the ECLS-K:2011 Kindergarten-Second Grade Data File and Electronic Codebook, Public Version (NCES 2017-285). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Vygotsky, L. (1987). Zone of proximal development. Mind in society: The development of higher psychological processes. 5291, 157.

Welsh, J. A., Nix, R. L., Blair, C., Bierman, K. L., & Nelson, K. E. (2010). The development of cognitive skills and gains in academic school readiness for children from low-income families. Journal of educational psychology, 102(1), 43.

Zelazo, P., Muller, U., Frye, D. & Marcovitch, S. (2003). The development of executive function in early childhood. Monographs of the Society for Research in Child Development, 68(3).

Zelazo, P. D., Anderson, J. E., Richler, J., Wallner-Allen, K., Beaumont, J. L., & Weintraub, S. (2013). II. NIH Toolbox Cognition Battery (CB): Measuring executive function and attention. Monographs of the Society for Research in Child Development, 78(4), 16-33.

Zelazo, P. (2015). Executive function: reflective, interactive reprocessing, complexity, and the developing brain. Developmental Review, 38, 55-68.

Zelazo, P. (2015). The dimensional change card sort (DCCS): a method of assessing executive functions in children. Nature Protocol, 1(1), 297-301.