A LONGITUDINAL INVESTIGATION OF LANGUAGE AND EXECUTIVE FUNCTION ON MATHEMATICS AND SCIENCE ACHIEVEMENT IN EARLY CHILDHOOD

by

Jocabed G. Marquez, B.S., M.S.

A dissertation submitted to the Graduate Council of
Texas State University in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
with a Major in School Improvement
August 2018

Committee Members:

Larry R. Price, Chair

Melissa A. Martinez, Co-Chair

Shetay N. Ashford

Marialena D. Rivera

This dissertation is dedicated to my daughter Sofia J. Garcia. You embody everything beautiful and innocent in this world. I love you-*I can do all things through Christ who gives me strength*.

ACKNOWLEDGEMENTS

First and foremost, I give praise and glory to my God Almighty who has always been by my side, without God I am nothing. I would like to thank the members of my dissertation committee: Dr. Larry R. Price, Dr. Melissa A. Martinez, Dr. Shetay N. Ashford, and Dr. Marialena Rivera. Special thanks to my mentor and dear friend, Dr. Larry Price, who for the past three years has met with me personally to discuss every detail of my study with patience and understanding. Many thanks to Dr. Melissa A. Martinez for her standard of excellence in research and scholarship. I greatly appreciate Dr. Shetay Ashford's view of STEM education and her unending support. Dr. Rivera's expertise and careful attention to detail was invaluable. Thank you to the faculty and staff from the School Improvement program who supported me in my journey: Dr. Miguel A. Guajardo, Dr. Duncan Waite, Dr. Clarena Larrota, Dr. Laura Rodriguez, Dr. Araceli Ortiz, Dr. Josyln Johnson, Dr. Kayon Murray, Dr. Chinedu Anumudu, Sara Garcia-Torres, Tammy Gonzales, Kayla Richardson, Tiffany Rodriguez, and Laura Shearer.

I want to thank all my family members: Samuel Garcia, Jr. Sofia J. Garcia, Socorro Guzman, Horacio Guzman, Guadalupe Marquez, Teresa Marquez, Maria del Socorro Guzman-Marquez, Eleazar Marquez, Eleazar Marquez Jr., Asael Marquez, Hadassa Ebisch, Marius Ebisch, Frida Ebisch, Julian Ebisch, Samuel Garcia Sr., Irma Garcia, Steven Garcia, Ashley Garza, and Natalie Garcia. I love you all very much-it is because of your prayers that I am here.

TABLE OF CONTENTS

Page

LIST OF TABLES

Table **Page**

LIST OF FIGURES

Figure **Page**

ABSTRACT

The United States faces high demand for science, technology, engineering, or mathematics (STEM) professionals and a scarce supply of individuals who pursue STEM careers, especially minority populations in the U.S with proficiency in a language other than English. The primary goal of this research was to determine the impact of use of Spanish in the home and direct cognitive assessments (executive function) on student achievement in mathematics and science during the fall of kindergarten, spring kindergarten, fall and spring of first grade, and fall of second grade. Parallel process longitudinal growth modeling was used to examine mathematics and science trajectories over time in a large cohort of students while simultaneously investigating tangential issues affecting change in achievement over time. Several analyses were employed in this study with the goals of: 1) Examining the growth of mathematics or science scores in isolation employing a univariate analysis model within the PPLGM, 2) Revealing the joint associations between growth factors capturing mathematics and science achievement employing an unconditional multivariate analysis and 3) Examining the effect of time-varying covariates as predictors of mathematics achievement scores at each year by employing a conditional multivariate analysis. Structural equation modeling (SEM) served as the analytic framework for conducting all analyses. This study used variables from the Early Childhood Longitudinal Study Cohort 2011.

I. Introduction

To achieve sustainable economic prosperity, nations depend on an educated, informed, and committed citizenry. To achieve economic security and stability, national administrations must invest in high-quality, expanded education for all its citizens. In today's globalized educational context, the rise of science, technology, engineering, and math (STEM) has gained significant traction and is viewed a critical portion of the formula designed to meet economic and social demands (Committee on Underrepresented Groups and the Expansion of the Science and Engineering Workforce Pipeline, et al., 2011). Globalization, competition and economic interests have driven decisions that impact educational reform. Nations from around the globe have prioritized and committed to a national reform agenda that call for drastically restructuring existing educational systems. For most countries, this means creating a STEM culture that promotes the conditions that nurture a workforce equipped with the skills to meet the demands of globalized society. Viewed as possessing both economic and moral imperatives, (National Science Teachers Association, 2017), countries are tasked with devising creative solutions to address the growing prominence of quality STEM educational opportunities for their respective populace.

This investment in education reform centers on dramatically improving and advancing a STEM driven agenda that creates the infrastructure to meet current and future demands. According to the President's Council of Advisors on Science and Technology (PCAST) (2010),

> "The success of the United States in the 21st century – its wealth and welfare – will depend on the ideas and skills of its population. These have always been the Nation's most important assets. As the world becomes increasingly technological, the value of these national assets will be determined in no small measure by the effectiveness of science, technology, engineering, and mathematics (STEM) education in the United States" (p.1).

For the United States to maintain the global leadership and competitiveness in science and technology, an investment in research and a strong innovative workforce are critical in achieving national goals. Yet, while science and engineering capabilities are as strong as ever, the dominance of the United States in the fields of STEM has lessened as the rest of the world has invested in and grown their education and research capacities (Institute of Medicine, National Academy of Sciences, and National Academy of Engineering, 2007).

Chapter One consists of the following themes: Background of the Problem, United States STEM Education in a Globalized Context, Underrepresented Students of Color in STEM, International Comparisons of K-12 Mathematics and Science Performance among U. S. Underrepresented Groups, Hispanic Students in STEM, Bilingual Hispanic Students and their Achievement in STEM, Executive Functioning and its Relation to Achievement, Statement of the Problem, Purpose of Study, Research Design, Research Questions, Epistemological Framework, Limitations, Significance of the Study, Definition of Terms, Organization of the Study, and Summary as seen in Figure 1 below.

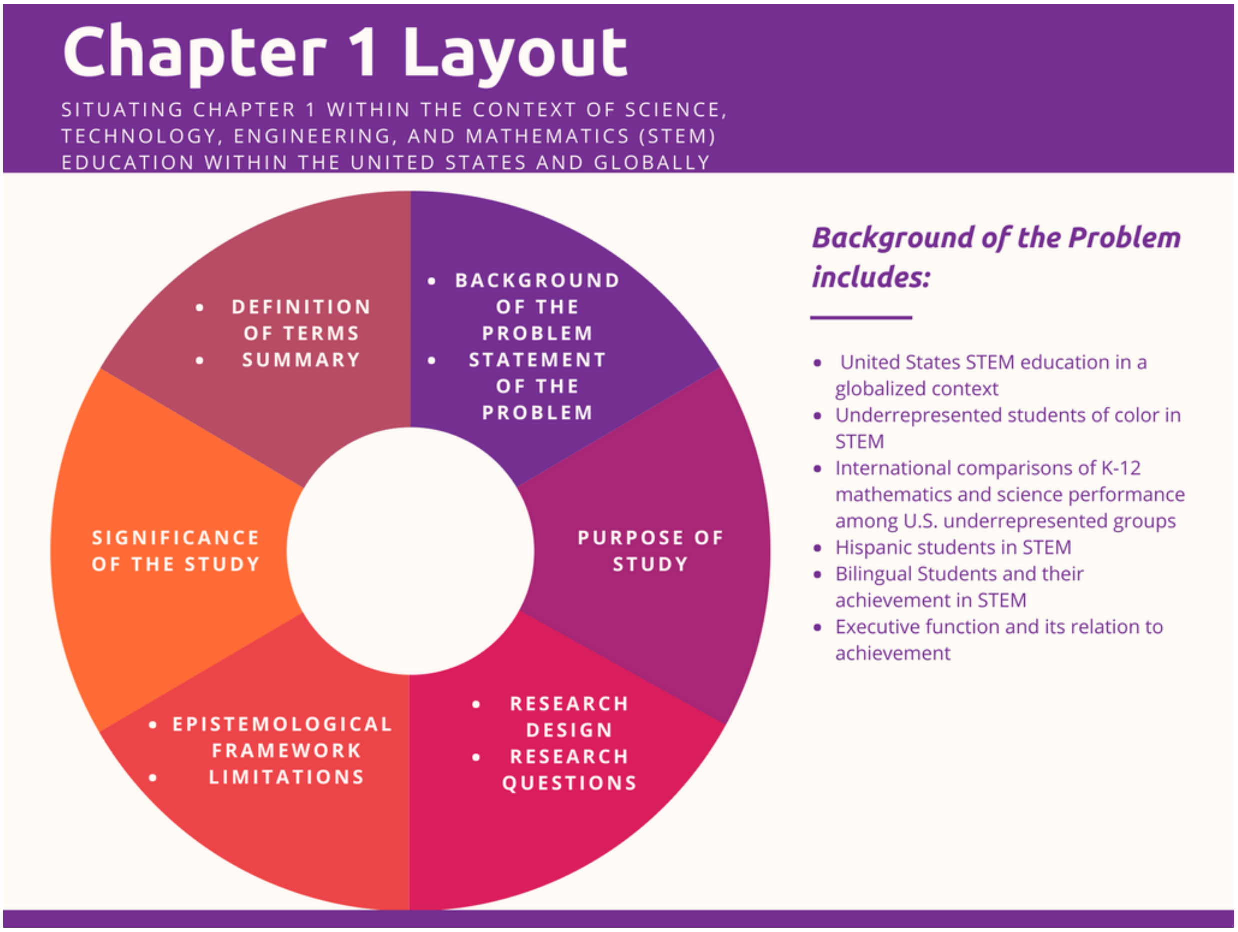

Figure 1. *Chapter One Layout.*

Background of the Problem

Since the 1970's overall educational attainment has stagnated in the United States, even as technological change and the return to higher education has increased. At the same time, most countries in Europe and several in Asia have surpassed the United States in educational attainment (Committee on Underrepresented Groups and the Expansion of the Science and Engineering Workforce Pipeline, et al., 2011). According to the United States President's Council of Advisors on Science and Technology executive report, despite the historical record of achievement in the United States, educational achievement at the elementary and secondary level is lower than in other nations in STEM education (President's Council of Advisors on Science and Technology, 2010). According to the advisors to the nation on science, engineering and medicine, international comparisons show many U. S. students fare poorly relative to their peers in other countries with regards to their understanding of science concepts (Duschl, Schweingruber, & Shouse, 2007). In the year 2000, the United States ranked 20 out of 24 countries in the percentage of 24-year-olds who had earned a first degree in the natural sciences or engineering with Finland, France, Taiwan, South Korea, and the United Kingdom surpassing the United States 10% benchmark percentage (Committee on Underrepresented Groups and the Expansion of the Science and Engineering Workforce Pipeline, et al., 2011).

Focusing specifically on K-12 education, international comparisons of student's performance in science and mathematics consistently places the United States in the middle or lower ranks on international assessments such as the Programme for International Student Assessment (PISA) and the Trends in International Mathematics and Science Study (TIMSS). In 2015, the Trends in International Mathematics and Science Study Scientific Literacy Assessment ranked Singapore first in eighth grade mathematics while the United States displayed a mean score of 510 (average is 500),18 points below the average, was ranked 12th out of 30 OECD countries (Provasnik, Malley, Stephens, Landeros, Perkins, & Tang, 2016) and was ranked 29th out of 40 industrialized nations (OECD, 2009). With a notable decline in rankings over the last three testing cycles, the U. S. is clearly underperforming while other countries are increasingly leading the way in the areas of science and mathematics. On the National Assessment of Educational Progress (NAEP), (Grigg, Lauko, Brockway, 2006) less than one third of U.S. eighth graders demonstrate proficiency in mathematics and science. Thirty six percent of Asian/Pacific Islander and twenty nine percent of White students scored at or above *Proficient*, while just 6 percent of Black, 8 percent of Hispanic, and 6 percent of American Indian/Alaska native students performed at that level. A closer examination of how these communities of color are underrepresented in STEM will follow. In this study Black, Hispanic, and American Indian/Alaska Native students will be referred to as underrepresented STEM students or underrepresented students of color in STEM. These same communities may be referred to as "minorities" only when this term is utilized in cited sources.

Underrepresented Students of Color in STEM

There are several critical issues for the U.S.' STEM infrastructure that remain unsettled to this day (Institute of Medicine, National Academy of Sciences, and National Academy of Engineering, 2007). America faces a demographic challenge with regard to its science and engineering workforce: Minorities are underrepresented in science and engineering yet are also the most rapidly growing sector of the population (Committee on Underrepresented Groups and the Expansion of the Science and Engineering Workforce

Pipeline, et al., 2011). Underrepresented minorities were largely and systematically excluded from mainstream educational opportunities through de jure and de facto segregation that continued from *Plessy v. Ferguson* in 1896 through the desegregation battles of the 1970s. This period of inclusion coincides with the period of increasing educational opportunity for White Americans. The period of inclusion for underrepresented minorities from the 1970's and beyond coincides with stagnation in both public educational investment and overall levels of educational attainment (Newfield, 2008). According to Newfield (2008), little progress has been made beyond marginally improving educational outcomes for minorities. Postsecondary attainment rates of underrepresented students is substantially lower as compared to White and Asian students. Native Americans, Alaska Natives, Hispanics, and African Americans must quintuple their proportions with a first degree in the STEM fields to achieve the global goal set at 10%. Based on the 2013 National Science Foundation report on Women, Minority, and Persons with Disability in STEM, among individuals who completed a bachelor's degree and attained a STEM career (N=5,069,000), 25% were female, 75% who were male, 2.4% were underrepresented minorities, and 69.3% were Whites. The same report found that US science and engineering workforce consisted of 0.4% American Indian, 16.4% Asian, 4.0% Black, 4.7% Hispanic, and 74.5% White, as compared to the total population of the United States with 1% American Indian, 4.4% Asian, 12.5% Black, 15% Hispanic, and 67.4% White (see Figure 1 below).

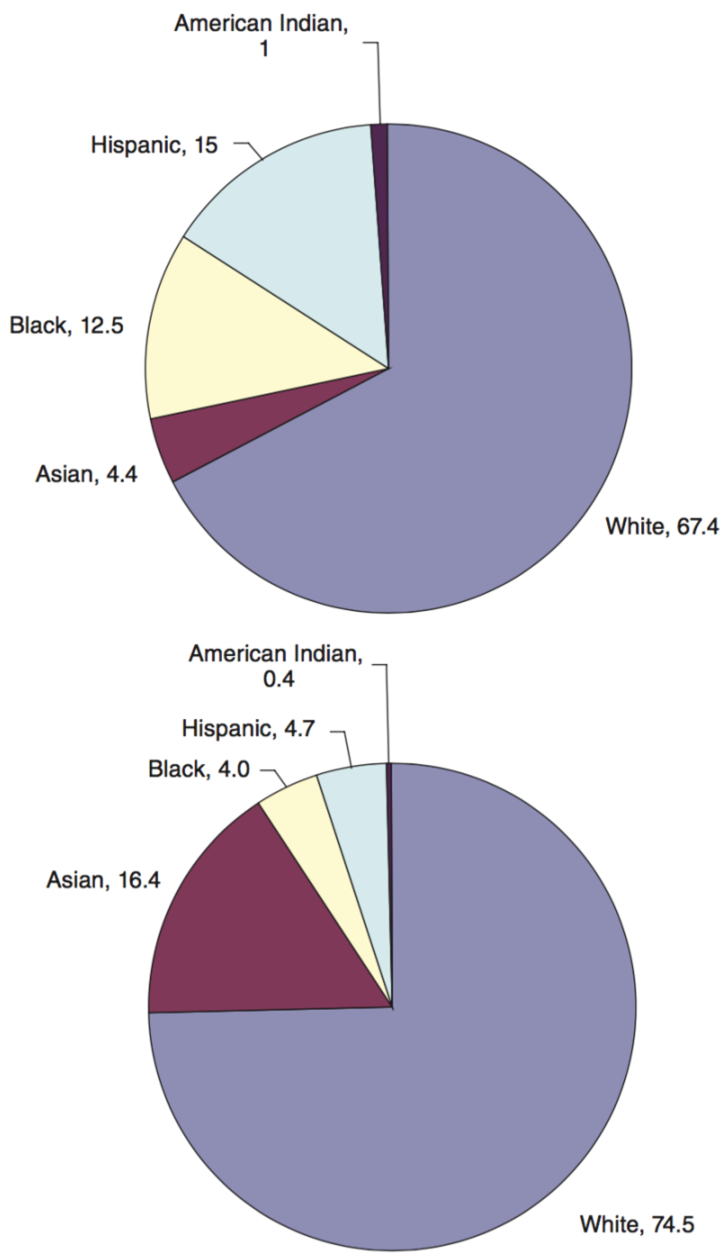

Figure 2. *US population and US Science and Engineering Workforce by 2006.* Source: National Science Foundation, Women, Minorities, and Persons with Disabilities in Science and Engineering, Tables A-2 and H-7. Numbers are percentages. http://www.nsf.gov/statistics.wmpd/

International Comparisons

The preliminary findings in the 2015 Trends in International Math and Science Study (Provasnik et al., 2016) documented that African American and Hispanic students were narrowing the gap in 4th grade mathematics, but Figure 2 for 8th grade mathematics a large gap remained. This figure only contains average scale scores and does not account for effect sizes that might give the reader a better insight on the discrepancies in scores.

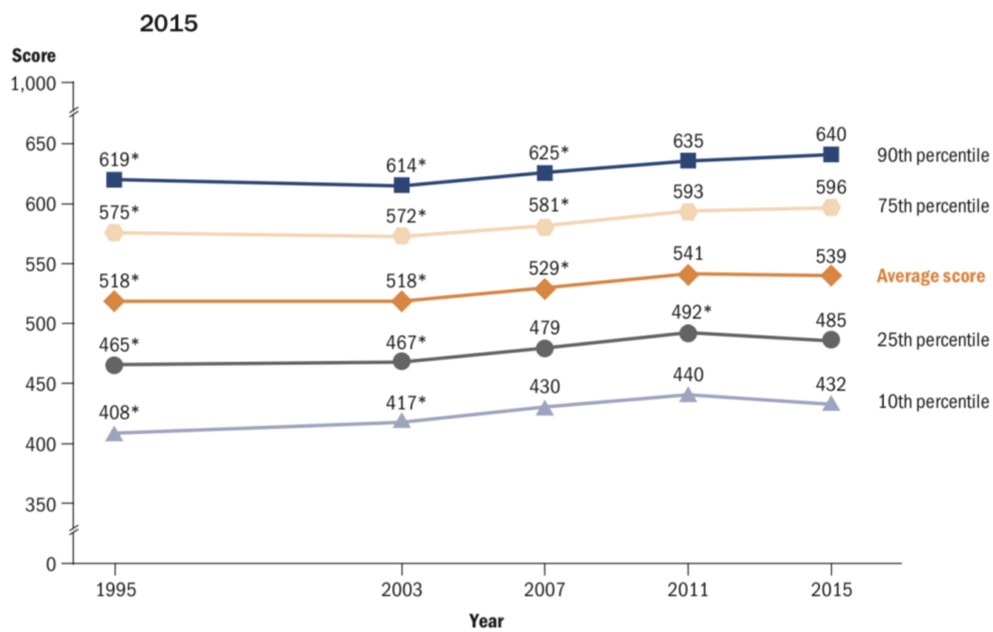

Figure 3. *Trends in U.S. 8th Grade Students' Average Mathematics Scores: 1995-2015*. Source: International Association for the Evaluation of Educational Achievement (IEA), Trends in International Mathematics and Science Study (TIMSS), 1995, 1999, 2003, 2007, 2011, and 2015.

The Education Trust conducted a secondary analysis of the TIMSS data and concluded that average mathematics and science scores for underrepresented minorities are below the national average and even less competitive globally (Miller, Malley, Burns,

2009) as compared to other Group of Eight (G8) countries-Canada, France, Germany, Italy, Japan, the Russian Federation, the United Kingdom-that are among the world's most economically developed. Figure 4 below plots the 4 major ethnic groups-Asian, White, Hispanic, and African American- tested on the international assessment, TIMSS. This graph depicts the discrepancies in scores amongst the four major ethnic groups compared to other developed and developing nations. Most alarming are the huge disparities identified in scores of both US Hispanic and African American students.

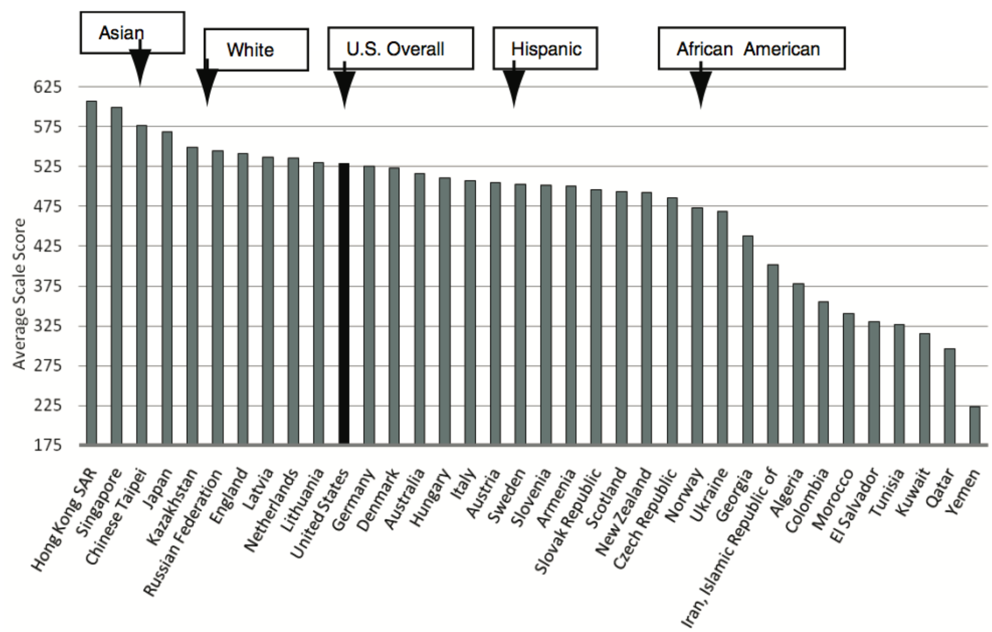

Figure 4. *TIMSS Grade 4 Math Racial/Ethnic Subgroup International Comparison.* SOURCE: Highlights from Trends in International Mathematics and Science Study (TIMSS) 2007, National Center for Education Statistics, U. S. Dept. of Education

Hispanic Students in STEM

Particularly noteworthy is the shortage in the number of Hispanic students entering the science, mathematics, engineering, and technology fields (Institute of

Medicine, National Academy of Sciences, and National Academy of Engineering, 2007). The Hispanic population in the United States has been steadily rising over the past century. In 2015, Hispanics made up 17.6% of the total population in the United States compared to 3.5% in 1960. In 2016, Hispanic individuals accounted for 18% of the total population of the United States and were the second largest racial group behind Whites (Flores, 2017). According to the Pew Research Center projections, it is estimated that the Hispanic population will comprise 24% of the total population in the United States by the 2065 (Flores, 2017). In 2006, Hispanic or Latino Americans comprised 15.0 percent of the U. S population and 17.8 percent of the college-age population, age 18-24. However, in 2005, they earned 7.9 percent of science and engineering bachelor's degrees and 6.2 percent of science and engineering master's degrees. In 2007, they earned 5.2 percent of science and engineering doctoral degrees awarded by U. S. institutions to U. S. citizens and permanent residents and 2.9 percent of S&E doctorates awarded to all recipients.

English-Language Learners and Their Achievement

A recent study of nationally-representative data found that 3 in 4 individuals of Hispanic descent born in the United States in the year 2001, were raised in homes where at least some Spanish is spoken (Lopez, Barrueco, & Miles, 2006). According to the U. S. Census (2010), 76% of Hispanics reported speaking primarily Spanish in their homes. According to recent studies on low-income preschoolers and their families, more than 25% of children in the United States use languages other than English in their home, with 92% of those children speaking Spanish (Moiduddin, Aikens, Tarullo, West, & Xue, 2012; Choi, Jeon, & Lippard, 2018). Such statistics indicated that most Hispanics are bilingual, or communicated in both English and Spanish, even if they are not being served in a bilingual program within an education setting.

Several terms are utilized in the literature to describe U. S. schoolchildren whose native language is a language other than English. The most common term is *language minority*, which is utilized to describe children whose native language is other than English and is applied to nonnative English speakers regardless of their current level of English proficiency. (Garcia, Arias, Harris Murri & Serna, 2010). Other potential terms that do not utilize “minority” are *non-native English speaker*, *Spanish dominant*, *linguistically diverse*, and *heritage language students*. More recently, the term dual language learner has been used to describe young language-learning children who are learning to speak their home language as well as at least one other language at the same time (Castro, Espinosa, & Paez, 2011). As noted in Espinosa (2007), other common terms common terms are *English language learner* (ELL), and *English Leaner* (EL). Other potential terms that do not utilize “minority” are *non-native English speaker*, *Spanish dominant*, *linguistically diverse*, and *heritage language students*. More recently, the term dual language learner has been used to describe young language-learning children who are learning to speak their home language as well as at least one other language at the same time (Castro, Espinosa, & Paez, 2011).

The percentage of public school students in the United States who were English language learners was higher in the Fall of 2015 with approximately 9.5 percent, or 4.8 million students compared to the Fall of 2000 with approximately 8.1 percent, or 3.8 million (McFarland, Hussar, Wang, Zhang, Wang, Rathbun, Barmer, Cataldi, & Bullock

Mann, 2018). In the Fall of 2015, Alaska, California, Colorado, Kansas, Nevada, New Mexico, Texas, and Washington were the states considered to have 10.0 percent or more of their total ELL population attending public schools (McFarland et al., 2018). In the Fall of 2015, there approximately 3.8 million Hispanic ELL students constituting 77.7 percent of the total ELL student enrollment followed by Asian ELL students which constituted roughly 10 percent of the overall ELL population overall. In addition, there were 29,500 White ELL students accounting for 6.1 percent of all ELL students and 178,00 Black ELL students accounting for 3.7 percent of ELL students. In each of the other racial/ethnic groups for which data were collected including Pacific Islanders, American Indians/Alaska Natives, and individuals of Two or more races, fewer than 40,000 students were identified as ELL's.

In fall of 2015, a greater percentage of public school students in lower grades than those in upper grades were ELL students. For example, 16.3 percent of kindergarteners were ELL students, compared to 8.2 percent of 6th graders and 6.6 percent of 8th graders. Explaining the pattern is driven in part by students identified as ELL's when they enter elementary school but obtain English language proficiency before reaching upper grades (Saunders, Marcelleti, 2013).

Klein, Bugarin, Beltranena, and McArthur (2004) found that 18 to 24-year-old English Learners who speak a language other than English at home and/or having varying English-speaking abilities, were less likely to be enrolled in college. Additionally, using the National Education Longitudinal Study of 1988 (NELS: 88), Kanno and Cromley (2013) study using the National Education Longitudinal Study of 1988 (NELS: 88), found that only one in eight English Language Learners (ELLs) attained a bachelor's degree compared to one in four English-proficient linguistic minorities and one in three native English speakers who earned a bachelor's degree. By the year 2030, approximately 40% of school aged children will be an English language learner or linguistic minority (Thomas & Collier, 2002).

STEM Education in Early Childhood

Since the implementation of No Child Left Behind Act Recent reports indicate very little time is dedicated to teaching STEM subjects in early childhood. Although the context of STEM education is commonly described as ranging from kindergarten to 12th grade, research on STEM education has focused primarily on upper elementary and secondary education settings (Merrill & Daugherty, 2010: Moorehead & Grillo, 2013). Teaching STEM in early childhood has received little attention. The National Research Council (2011) has emphasized the need to include kindergarten to third grade in advancing K-12 STEM education.

As evidenced by a growing body of research, early STEM experiences (defined as preschool to third grade) play an important role in enhancing children's knowledge, skills, and dispositions needed to fulfill next-generation employment (Park, Dimitrov, Patterson, & Park, 2017). Preparing students for an economy demanding innovative solutions to complex problems should begin during early childhood (Aronin & Floyd, 2013; Chesloff, 2013; DeJarnette, 2012; New, 1999). For example, Chesloff (2013) argued that STEM education should start in early childhood since "concepts at the heart of STEM-curiosity creativity, collaboration, critical thinking-are in demand" (p. 27). Although research has indicated the need to begin STEM in early childhood, the idea of teaching STEM to young children ages 3 to 8 remains elusive and marginalized by teachers and administrators in schools (Parette, Quesenberry, & Blum, 2010).

Two studies have found that in a typical school day at the Pre-Kinder to third grade level classrooms, language arts accounted for 89 minutes of instruction, math accounted for 54 minutes, and science accounted for only 19 minutes, suggesting there is a small likelihood a child will receive any form of exposure to STEM activities (Horizon Research, 2013; National Research Council, 2011). Such exposure to science and mathematics is especially necessary during the early elementary school years when children's skills are more malleable (Bornstein, Hahn, Putnick, & Suwalsky, 2014). Kindergarten attendance, having grown exponentially over the last several decades (Davis & Bauman, 2013), constitutes a critical development period for shaping children's long-term life success (Duncan et al, 2007; Jones, Greenberg & Crowley, 2015). The quality of children's learning environments prior to age six has an influence on later academic success (e.g. Campbell, Pungello, Miller-Johnson, Burchinal & Ramey, 2001: Hadzigeorgiou, 2002). With the rise in the number of children attending kindergarten, coupled with the heightened emphasis on children's early learning (Bassok, Latham, & Rorem, 2016) it is essential to address STEM instruction, executive functioning skills, and language in the early on. In the Fall of 2015, a greater percentage of ELL students in public schools were concentrated in kindergarten, first grade, second grade, and third grade.

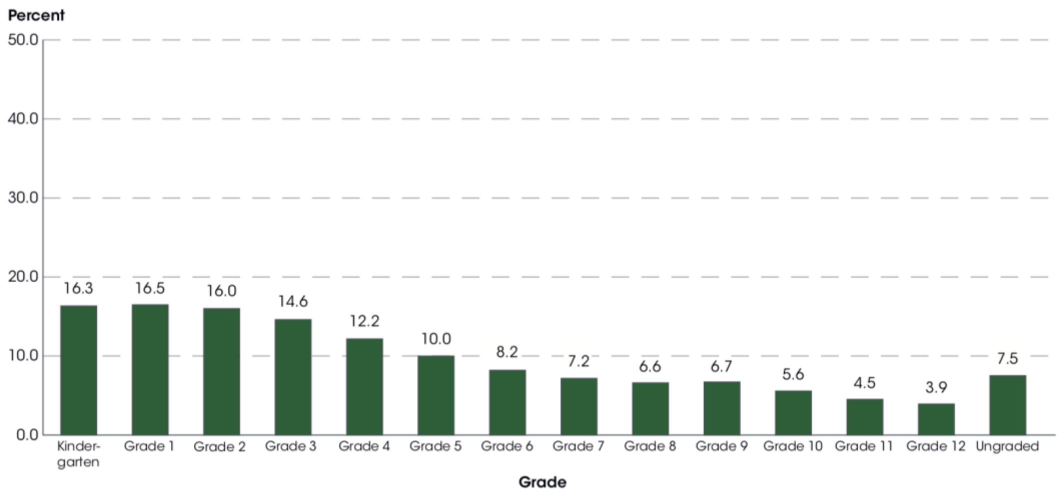

Figure 5. *Percentage of K-12 English Language Learners Fall 2015.* Source: US Department of Education National Center for Education Statistics

Source: US Department of Education National Center for Education Statistics

Approximately sixteen percent of kindergarteners were ELL students compared to only 3.9 percent of 12th grade ELL students. This observed pattern explains the development of English proficiency throughout the progression of schooling from elementary to secondary (Saunders & Marcelletii, 2013). Given the high percentage of ELL students in the early grades in public schools, this study focuses on students in the early grades and their achievement in mathematics and science while examining measures of executive functioning.

Executive Function and its Relation to Achievement

A number of studies have shown a developmental link between executive function and performance in mathematics and science, especially in school-aged children (Agostino, Johnson, & Pascual-Leone, 2010; Harvey & Miller, 2017; Van der Ven, Kroesbergen, Boom, & Leseman, 2012). Such studies are crucial in identifying cognitive precursors of mathematics achievement before school entry and contribute to the development of interventions that may play a critical role in enhancing children's learning of early math concepts (Veterbori, Usai, Traverso, De Franchis, 2015). Executive functions are concerned with the regulatory processes that allow for the initiation, modulation, and inhibition of ongoing mental attention necessary for task performance (Dempster, 1992; Dennis, 1991). Executive function skills make it possible for persons to sustain attention, keep goals and information in mind, refrain from responding immediately, hinder distraction, tolerate frustration, acknowledge behavior consequences, and plan for future events (Zelaso, Blair, & Willoughby, 2017). Similarly, children for whom English is not their primary language, and therefore not the language of their home environment, exhibit low academic performance not only in reading and writing, but also in math and science (Guglielmi, 2012; National Center for Education Statistics, 2009, 2011). To better understand this phenomenon, this study leverages a nationally representative data set with direct assessment of elementary school-aged children's executive function skills to model growth in mathematics and science achievement at over time. This study specifically focuses on two variables of EF-working memory (WM) and cognitive shifting also known as cognitive flexibility.

Statement of the Problem

The demand for STEM graduates in STEM fields continues to grow at a relatively rapid rate. According to the National Science Foundation (2010), the employment rate in science and engineering fields rose an average of 1.3% in all occupations, and this estimated growth rate is consistent with long-term national trends (US Department of

Labor, 2007). By 2018, nine of the fastest growing occupations that require at least a bachelor's degree will depend on significant math or science training, and many science and engineering occupations are predicted to grow faster than the average rate for all occupations (National Science Board, 2010; US Census Bureau of Labor Statistics, 2014). Crucial to the discussion on broadening STEM participation is the underrepresentation of racial minorities, women, and students of low socioeconomic status (Anderson & Kim, 2006; Herrera & Hurtado, 2011; National Science Foundation, 2010; 2015; Schultz et al., 2011). Addressing issues of equity must be a top priority for policy makers and other educational leaders that make pivotal decisions that impact the lives of children across the country.

Although a world-leading STEM workforce is necessary, according to the Presidential Advisory Commission on Educational Excellence for Hispanic Americans (2003), the nation is "losing Hispanic American students all along the education continuum" (p.1): In the year 2005, one in every three Hispanics had not completed high school, and among those who completed high school, only 53% enrolled in postsecondary education after graduation, compared to 66% of non-Hispanic Whites (Lindholm-Leary & Borsato, 2005). As the demographic trends shift toward a higher representation of Hispanics in the United States (24% of the U.S. population by 2050, Pew Research Center), the academic attainment of Hispanic students becomes crucial.

The shortage of Hispanic students entering the science, mathematics, engineering, and technology fields is striking. Data from the National Assessment of Educational Progress (NAEP) consistently reveal mathematics-related achievement discrepancies related to income and race (National Mathematics Advisory Panel, 2008: Roberts & Bryant, 2011). The National Center for Education Statistics (NCES) and the Institute for Education Sciences (IES) in their latest publication, Digest of Education Statistics (2015), found that US Hispanic 4th grade student scores on the mathematics NAEP assessment were 18 points lower than White students. At grade 8, mathematics scores on the NAEP demonstrated a wider achievement gap in scores with a difference of 22 points between Hispanic and White students (Digest of Education Statistics, 2015). The most recent publication also examined Science NAEP scores and a 27-point gap between Hispanic and White student scores (Digest of Education, 2015). Research findings suggest that Hispanic students as group, are provided fewer opportunities in schools to acquire high-order skills in mathematics and science compared to their White counterparts (Clark, 1999, Jensen, 2007; Lindholm-Leary & Borsato, 2001; Lindholm-Leary & Borsato, 2005; Marian, Shook, & Schroeder, 2013; Strutchens & Silver, 2000; Tochon, 2009). One of the opportunities not given to Hispanic students is to develop their mathematics and science skills early on by building on their language development. Hispanic students are not given the opportunity to develop high-order thinking skills in mathematics and science.

Purpose of the Study

The focus of this research is to determine the impact of use of Spanish in the home and direct cognitive assessments (executive function) on student achievement in mathematics and science during the fall of kindergarten, spring kindergarten, fall first grade, spring first grade, and fall of second grade. According to the literature, the highest concentration of English learners in U.S.' schools are in the early grades and they tend to develop oral and academic English proficiency throughout the early grades (Halle, Hair, Wandner, McNaman, & Chien, 2012). This study utilizes variables from the Early Childhood Longitudinal Study Cohort 2011. The Early Childhood Longitudinal Study is sponsored by the National Center for Education Statistics (NCES), within the U.S. Department of Education's Institute of Education Sciences, to provide detailed information on the school achievement and experiences of students throughout their elementary school years (McCarroll, Flanagan, & Potter, 2016). The students participating in the ECLS-K: 2011 are assessed longitudinally from kindergarten (the 2010-2011 school year) through the spring of 2016, when most are expected to be in fifth grade. The design of the ECLS-K:2011 and its survey instruments is guided by a conceptual framework of children's development and learning that emphasizes the interaction among the various environments in which children live and learn and the resources within those environments to which children have access to (Tourangeau, Nord, Wallner-Allen, Vaden-Kiernan, Blaker & Najarian, 2017).

Figure 6 below summarizes the present study's conceptual framework and describes how executive functioning and achievement are linked to math and science outcomes. The figure also demonstrates how improved executive function among bilingual speakers and how it relates to mathematics and science achievement. All three circles combine and the current study is seen among all the themes.

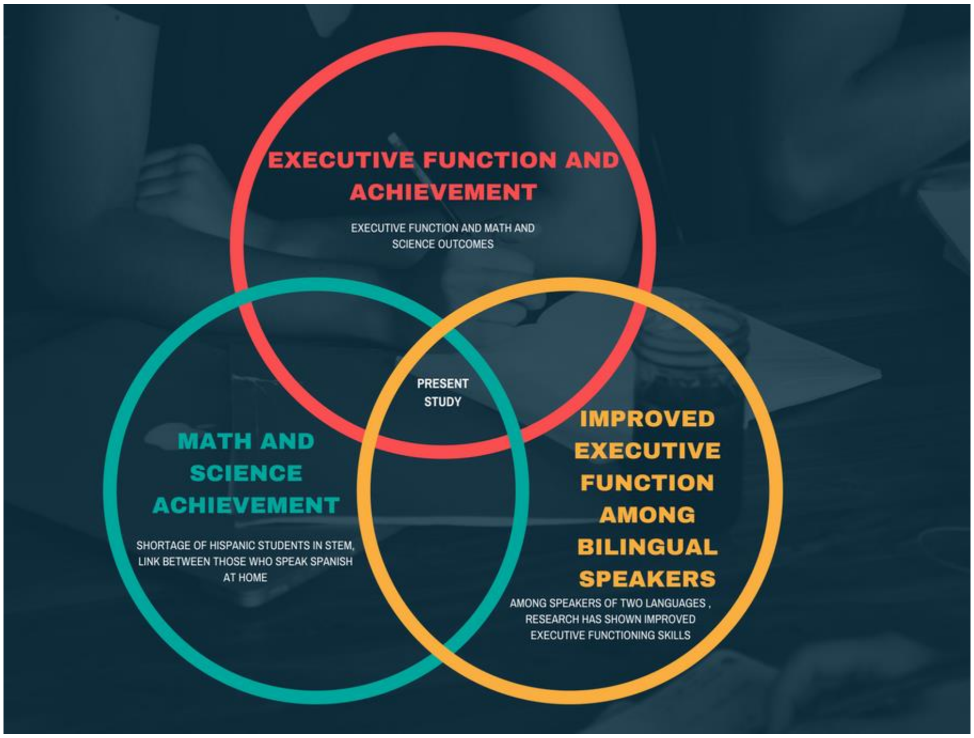

Figure 6. *Conceptual Framework of Present Study*

Research Design

In this study, parallel process longitudinal growth modeling was utilized to simultaneously study mathematics and science trajectories over time in a large cohort of students and to examine several critical issues related to potential changes in achievement over time. Using the PPLGM, the outcomes are threefold based on three types of analyses each serving a different purpose: 1) To examine the growth of mathematics or science in isolation employing a univariate analysis model within the PPLGM, 2) To reveal the joint associations between growth factors of mathematics and science employing an unconditional multivariate analysis and 3) To examine the effect of time-varying covariates as predictors of mathematics achievement scores at each year by employing a conditional multivariate analysis. Structural equation modeling (SEM) serves as the analytic framework for conducting our analyses. This study utilizes variables from the Early Childhood Longitudinal Study Cohort 2011. The Early Childhood Longitudinal Study is sponsored by the National Center for Education Statistics (NCES), within the U.S. Department of Education's Institute of Education Sciences, to provide detailed information on the school achievement and experiences of students throughout their elementary school years (McCarroll, Flanagan, & Potter, 2016). The students participating in the ECLS-K: 2011 are assessed longitudinally from kindergarten (the 2010-2011 school year) through the spring of 2016, when most are expected to be in fifth grade.

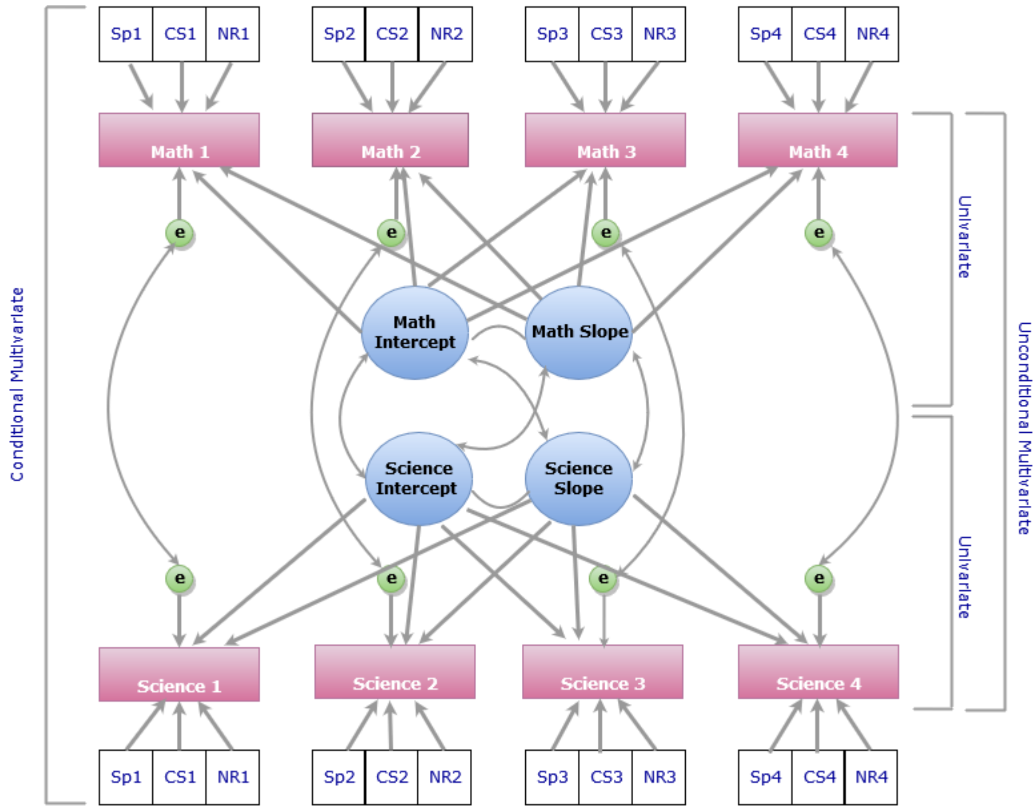

Figure 7. *Parallel Process Growth Curve Model of Present Study***.** The latent factors are shown as ovals (Math Intercept, Math Slope, Science Intercept, and Science Slope) on four different time points. The covariates correspond to Sp1 (Spanish spoken at home), CS2 (Card Sort Combined Score), and NR (Numbers Reversed W Score) are shown as rectangles on four different time points.

Research Questions

1. Is the initial level (*intercept* a.k.a. the "mean") for math and science significantly different for students in the non-Spanish speaking home environment as compared to students who speak Spanish dominantly in the home?

2. Is there significant *variance* of the intercept (a.k.a. the "mean") in math and science? If so, is it higher for students in the non-Spanish speaking home environment as compared to students who speak Spanish dominantly in the home?
3. Is the slope (rate of change) in math and science significantly different for students in the non-Spanish speaking home environment as compared to students who speak Spanish dominantly in the home?
4. Is the variance in the slope (variation around the rate of change) in math and science significantly different for students in the non-Spanish speaking home environment as compared to students who speak Spanish dominantly in the home?
5. Is there a significant correlation between the intercept and slope for math scores? If so, is it positive or negative?
6. Is there a significant correlation between the intercept and slope for science scores? If so, is it positive or negative?
7. Is there a significant correlation between the intercepts and slopes for math and science scores in the multivariate PPLGCM? If so, is it positive or negative?
8. Does executive functioning serve as a significant covariate for math and science score performance in the PPLGCM? If so, how much variance in math or science achievement does it explain?
9. Does a PPLGCM without executive functioning covariates fit statistically better or worse than the model with covariates?

Examination of the research questions above implies a need for a growth curve model that adequately accounts for change over time in mathematics and science achievement scores. Using growth curve analysis, it is also possible to evaluate to what extent executive functioning deficits are uniquely associated with kindergarten, first grade, second grade, and third grade risk of later experiencing science and mathematics difficulties. To this end, the analytic sample included a population-based and multi-year longitudinal sample of U.S. schoolchildren, including two measures of two types of executive function (working memory, cognitive flexibility). The overarching goal is to use theoretically driven covariates-my study includes Spanish of the home and executive functioning to explain the variation in initial levels (intercepts) and slopes (over four time points) across students.

Epistemological Framework

The epistemological framework guiding my study emerged as a reaction of educational researchers to the limitations of positivism as a paradigm (Abdul Hameed, Sanaullah, & Asif Ali 2017). Educational researchers contended with the limitations of positivism for social sciences' research and combined positivism with interpretivism to form a new paradigm named post-positivism (Petter & Gallivan, 2004; Deluca, Gallivan, & Kock, 2008). The post-positivist critical realist firmly believes that "the goal of science is to hold steadfastly to the goal of getting it right about reality, even though we can never achieve that goal" (Trochim, 2006). Since most observation and measurement is fallible and all theory is revisable, the post-positivist emphasizes the importance of multiple measures and observations, each of which may possess different types of error (Trochim, 2006). Post-positivists reject the idea that any individual can see the world perfectly as it really is therefore triangulation across multiple sources is necessary (Trochim, 2006). The current study applies post-positivism by not employing strict hypothesis testing and recognizing that social science research is not entirely without bias and subjectivity.

Study Limitations

The variables and data utilized in the present study originates from a publicly available dataset, ECLS-K: 2011 of the National Center for Education Statistics. Some of the limitations in my study include the analysis of a one national dataset containing no identifying information on the participants, therefore, it is not possible to interview participants for additional insights into my study. The scope of my study is limited to certain variables on a national dataset and therefore, does not account for external socio-political factors that have historically been documented in the research to affect student outcomes and achievement.

The current study is limited to those participants having a mathematics and science IRT theta score in the kindergarten, first and second grade, as well as a score on the variables used as covariates (executive function and Spanish of the home). The current study is limited to those participants who indicated Spanish is the language of the home. The current study is limited to those students attending public schools in city/suburb locale.

Significance of the Study

Existing literature focuses on distinct elements of STEM education but has not focused on assessing students in mathematics and science longitudinally through parallel process growth curve modeling, particularly with measures of executive functioning and Spanish in the home. This study is the first to follow children longitudinally in their early childhood trajectory examining the change over time in their mathematics and science achievement and whether Spanish and executive function significantly affect their achievement.

Little is known about language minorities' executive functioning and their achievement in science and mathematics. There has been some research that distinguish groups who purse STEM fields compared to groups who do not pursue STEM fields by individual characteristics such as students' math and science attitudes, self-efficacy in math and science, gender, race/ethnicity, and structural characteristics including socioeconomic status, immigrant generation status, prior achievement in math and science, tracking, course taking patterns, and extracurricular involvement. However, there has been no research that differentiates groups by language proficiency (i.e. English speakers and non-native English speakers), specifically in the early grades. The gap in the literature informs the current proposed study. The proposed research study will broaden understanding of the underlying mechanisms related to mathematics and science achievement, inform educational research, and impact classroom practice. The contents and the intended outcomes of this study will serve to inform decisions that affect educational policy, research, and practice and facilitates systemic change within our American public education system.

Large Comprehensive Assessments

Large comprehensive assessment data sets provide the basis for important secondary analyses that have implications regarding student, school, and cultural attributes far beyond the league tables. There are two main advantages for using the database of large-scale surveys for a secondary analysis: The scrupulous sampling design assures the sample is representative of the whole population, and the quality of the instrument is confirmed by pilot studies and by content and methodology experts.

Definition of Terms

ELL/Language Minority/Non-Native Speaker. Individuals whose native language is a l language other than English (Garcia, Arias, Harris Murri & Serna, 2010)

Cognitive Flexibility: The mental ability to switch between thinking about two different concepts, and the think about multiple concepts simultaneously.

CS: Dimensional Card Sort Game. Measure of executive functioning, specifically cognitive shifting/cognitive flexibility.

ECLS-K:2011: Early childhood longitudinal study Cohort 2011, children will be assessed from Kinder in 2011 all through 5th grade in 2016. Data collected is representative of 18,174 children enrolled in 968 schools. Data for the ECLS-K:2011 are released in both a restricted use and a public-use version. It is a multisource, multimethod study that focuses on children's early school experiences. It includes interviews with parents, selfadministered questionnaires completed by teachers and school administrators, and one on-one assessments of children.

NR: Numbers Reversed Game. Measure of executive functioning, specifically working memory.

Sp 1-4: Spanish of the home dichotomous variable assessed at Time 1, Time 2, Time 3, Time 4

STEM: Science, Technology, Engineering, and Mathematics (NSF, 2011).

Working Memory: The system responsible for the transient holding and processing of new and already-stored information, and is an important process for reasoning, comprehension, learning, and memory updating.

Organization of the study

The research presented in this document proceeds as follows. First, Chapter Two provides a review of the literature chronological literature and reviews of executive function and Spanish as they relate to mathematics and science achievement. Chapter Three provides a description of the methodology used or this study, research questions, and associated hypotheses. Chapter Four summarizes the analysis of data accessed from the Early Childhood Longitudinal Study Cohort 2011 as described in the methodology. Chapter Five provides discussion of results and implications for school leaders and overall school improvement based on the data presented.

Summary

For the United States to maintain the global leadership and competitiveness in science and technology, an investment in research and a strong innovative workforce are critical in achieving national goals. Yet, while science and engineering capabilities are as strong as ever, the dominance of the United States in the fields of STEM has lessened as the rest of the world has invested in and grown their education and research capacities (Institute of Medicine, National Academy of Sciences, and National Academy of Engineering, 2007). Critical is the shortage in the number of Hispanic students entering the science, mathematics, engineering, and technology fields (Institute of Medicine, National Academy of Sciences, and National Academy of Engineering, 2007). A number of studies have shown a developmental link between executive function and performance in mathematics and science, especially in school-aged children. Similarly, children for whom English is not their primary language, and therefore not the language of their home environment, exhibit low academic performance not only in reading and writing, but also in math and science. The existence of co-occurring factors and their relationship to how students' mathematics and science achievement scores change over time presents a substantial challenge to providing effective student-centered interventions. To increase understanding about the complex relations between student-level covariates and achievement scores, we introduce a parallel process latent growth model (PPLGM) with covariates. Structural equation modeling serves as the analytic framework for the PPLGM.

II. REVIEW OF LITERATURE

Introduction

The ability to meet the challenges and achieve the opportunities in a globalized context depends in large measure in our nation's science and engineering workforce. The importance of science and engineering to the United States has been documented in a series of reports over more than half a century, from Vannevar Bush's *Science, The Endless Frontier* (1945) to Deborah Shapley and Rustum Roy's *Lost at the Frontier* (1985) to the National Academies' Rising Above the Gathering Storm (2007). *Rising Above the Gathering Storm* argued that the United States is at a crossroads and that in order for it maintain its global leadership and competitiveness in science and technology that are critical to achieving national goals today, an investment in research, increased encouragement in innovation, and grow a talented and innovative science and technology workforce (IOM, NAS, NAE, 2007). According to the *Gathering Storm* (2007), in order for the United Sates to maintain the global leadership and competitiveness in science and technology that are critical in achieving the advancement of our nation, there are several key components that must align in order to achieve such goals. Citing the need to develop a strong and diverse science and engineering workforce, United States senators Edward Kennedy, Barbara Mikulski, Patty Murray, and Hillary Clinton requested in a study of underrepresented minority participation in science and engineering in November 2006. The U.S. Congress later included this request as a mandate in the 2007 America COMPETES Act, charging the study committee to "explore the role of diversity in the STEM workforce and its value in keeping America innovative and competitive, analyze the rate of change and the challenges the nation currently faces in developing a strong and diverse workforce, and identify best practices and the characteristics of these practices that make them effective and sustainable" (Committee on Underrepresented Groups and the Expansion of the Science and Engineering Workforce Pipeline, et al., 2011, p. 2).

The study committee then identified three reasons for the participation of underrepresented groups in science and engineering and how those reasons play a central role in sustaining the United States' research and innovation capacity:

1. *Sources for the future science and engineering workforce are uncertain*: For many years the United States relied on a STEM workforce that was predominantly and overwhelmingly Asian and White. Non-US citizens, particularly from India and China, have accounted for almost all growth in STEM doctorate awards. Relying on non-U.S citizens for science and engineering workforces is very uncertain (Committee on Underrepresented Groups and the Expansion of the Science and Engineering Workforce Pipeline, et al., 2011).
2. *The demographics of the United States are shifting dramatically*: Those groups that are most underrepresented in the science and engineering workforces are also the fastest growing in the general population (Committee on Underrepresented Groups and the Expansion of the Science and Engineering Workforce Pipeline, et al., 2011).

3. *Diversity is an asset*: Increasing the participation of underrepresented groups in science and engineering contributes to the talent pool, enhancing innovation, and improving the nation's global economic leadership (Committee on Underrepresented Groups and the Expansion of the Science and Engineering Workforce Pipeline, et al., 2011).

In line with reason three above, of particular concern in the discussion on broadening STEM participation is the underrepresentation of racial minorities, women, and students of low socioeconomic status (Anderson & Kim, 2006; Herrera & Hurtado, 2011; National Science Foundation, 2010; 2015; Schultz et al., 2011). Addressing issues of equity must be a top priority for policy makers and other educational leaders that make pivotal decisions that impact the lives of children across the country.

Condition of Hispanics in the United States

The Hispanic population in the United States has been steadily rising over the past century. In 2015, Hispanics made up 17.6% of the total population in the United States compared to 3.5% in 1960. In 2016, Hispanic individuals accounted for 18% of the total population of the United States and were the second largest racial group behind Whites (Flores, 2017). According to the Pew Research Center projections, it is estimated that the Hispanic population will comprise 24% of the total population in the United States by the 2065 (Flores, 2017). In 2014, eight states had Hispanic/Latino populations of at least 1 million: California, Texas, Florida, New York, Illinois, Arizona, New Jersey, and Colorado (Stepler & Lopez, 2016). The growth of the Hispanic population accounted for over half (54%) of the total United States population growth from 2000 to 2014. The Hispanic/Latino population in the United States has now reached 50 million and has been the principal drive of U.S. demographic growth, accounting for half of the national population growth since 2000. According to the Pew Research Center, “Hispanics are also the nation’s second-fastest growing ethnic group, with a 2.0% growth rate between 2015 and 2016 compared with a 3.0% rate for Asians.

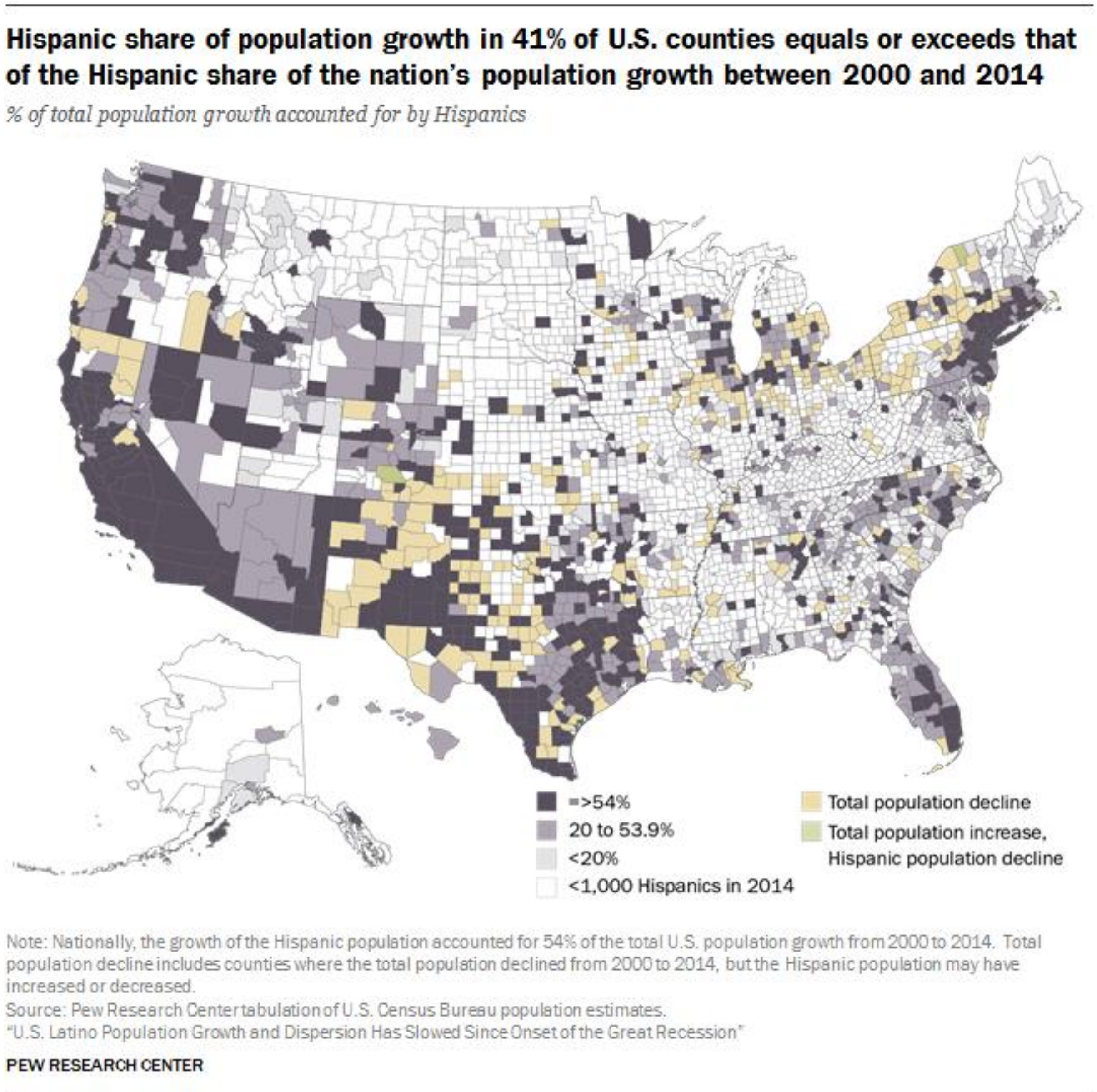

Figure 8. *Hispanic Share of Population Growth Between 2000 and 2014*. Source: Pew Research Center by R. Stepler & M.H. Lopez. Retrieved from http://www.pewhispanic.org/2016/09/08/3-where-hispanic-population-growth-is-driving-general-population-growth/Pew Research Center by R. Stepler & M.H. Lopez. Retrieved

from http://www.pewhispanic.org/2016/09/08/3-where-hispanic-population-growth-is-driving-general-population-growth/

In 2012, Whites represented 50% of the U.S population under the age of five, while Latinos comprised 26%, African Americans 13%, and Asians 7% (NCES, Digest of Education Statistics, 2012). For purposes of this study, the population of interest includes children in the age ranges of five, six, seven, eight, and nine years of age.

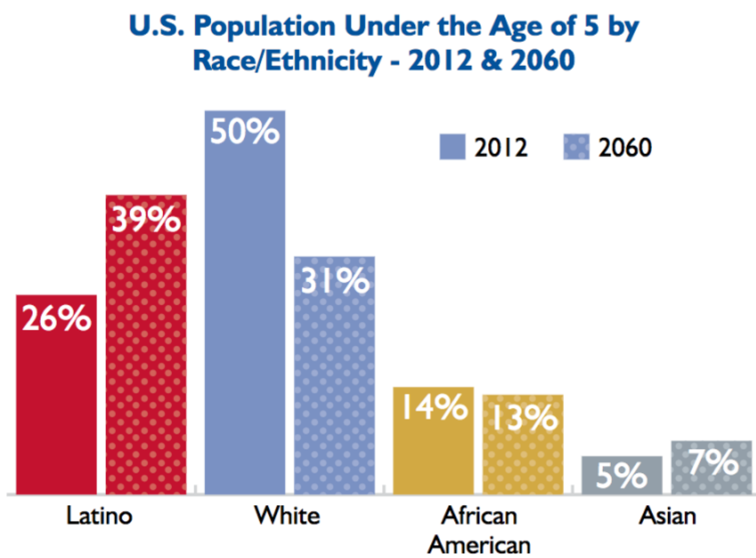

Figure 9. *United States Population Under Age 5 by Race/Ethnicity*. Adapted from U.S. Census Bureau, Population Division, Projected Population by Single Year of Age, Sex, Race, and Hispanic Origin for the United States: 2012 to 2060.

Linguistic Diversity

A recent study based on a nationally-representative data found that 3 in 4 individuals of Hispanic descent born in the United States in the year 2001, were raised in homes where at least some Spanish is spoken (Lopez, Barrueco, & Miles, 2006).

According to the U. S. Census (2010), 76% of Hispanics, the fastest growing ethnic group in the United States, reported speaking primarily Spanish in their homes. As of 2010, 20% of the United States school population entered the public school system speaking Spanish as their native language with limited proficiency in English (Esposito & Baker-Ward, 2013). In early education, the number of kindergarteners beginning public school in the United States who speak Spanish as their native language, known as Spanish Speaking Kindergartners (SSK) continues to increase (Jensen, 2007; Hernandez, 2006). According to recent studies on low-income preschoolers and their families, more than 25% of children in the United States use languages other than English in their home, with 92% of those children speaking Spanish (Choi, Jeon, & Lippard, 2018; Moiduddin, Aikens, Tarullo, West, & Xue, 2012). The next most spoken non-English languages are Chinese (with 2.8 million speakers), Hindi, Urdu or other Indian languages (2.2 million), French or French Creole (2.1 million) and Tagalog (1.7 million) (Gonzalez-Barrera & Lopez, 2013). Additionally, Spanish was the home language of 3.7 million ELL students in the Fall of 2015, representing 7.6 percent of all public K-12 students (McFarland et al., 2018). Arabic, Chinese, and Vietnamese were the next most common home languages spoken by approximately 114,400; 101.300; and 81,200 students respectively. Somali, Hmong, Russian, Haitian, Tagalog, Korean, Nepali, and Karen were the next most commonly reported home languages of ELL's in the Fall of 2015. English was the 5th most commonly reported home language for ELL students (80,300). Please see figure 10 below. It is important to note, not all Spanish speakers are Hispanic (Gonzalez-Barrera & Lopez, 2013). When it comes to English proficiency, 80% of non-Hispanics who speak

Spanish at home say they speak English “very well”, 11% say they speak English “well” and 9 % say the speak English “not well” or do not speak English (Gonzalez-Barrera & Lopez, 2013).

Home language	Number of ELL students	Percentage distribution of ELL students[1]	Number of ELL students as a percent of total enrollment
Spanish, Castilian	3,741,066	77.1	7.6
Arabic	114,371	2.4	0.2
Chinese	101,347	2.1	0.2
Vietnamese	81,157	1.7	0.2
English[2]	80,333	1.7	0.2
Somali	34,813	0.7	0.1
Hmong	34,813	0.7	0.1
Russian	33,057	0.7	0.1
Haitian, Haitian Creole	30,231	0.6	0.1
Tagalog	27,277	0.6	0.1
Korean	27,268	0.6	0.1

Figure 10. *English Language Learner Languages Fall 2015*. Source: U.S Department of Education, National Center for Education Statistics, ED*Facts* file 141, Data Group 678: and Common Core of Data (CCD).

Home language and Academic Achievement

Several studies have focused on the reading and mathematics achievement patterns from kindergarten through third grade of students living in homes categorized as “primarily Spanish” or “Spanish only” (Gormley, Gayer, Phillips, & Dawson, 2005; Reardon & Galindo, 2006).

The aforementioned studies have shown students living in homes categorized as “primarily Spanish” or “Spanish only” lagged behind their English-only Hispanic peers, Whites, and Asian Americans in all subjects at the beginning of and throughout early education (Gormley, Gayer, Phillips, & Dawson, 2005; Reardon & Galindo, 2006). Research on the relationship between language use in the home and ELL’s literacy development in their first or second language indicate a positive, although modest, relationship between the home use of a language and literacy achievement in that language, and conversely, a negative, very modest relationship between home use of a language and achievement in another language (August, 2006; Howard et al., 2014).

Several systematic literature reviews on language, literacy, and academic development by English learners reveal the relationships between English learners’ home language competencies and their academic development (August & Shanahan, 2006: Genesee 2002, 2004, 2007; Genesee, Boivin, & Nicoladis, 1996; Genesee & Gandara, 1999; Genesee & Geva, 2006; Genesee, Lindholm-Leary, Saunders, & Christian, 2005; Genesee, Lindholm-Leary, Saunders, & Christian, 2006; Genesee, Paradis, & Crago, 2004; Greene, 1998; MacSwan, Thomson, deKlerk, & McAlister, 2007; Roldstad, Mahoney, & Glass, 2005; Slavin and Cheung, 2005; Willig 1985). English learners with advanced levels of competence in certain aspects of the home language demonstrate superior achievement in English literacy compared with English learners who lack or have lower levels of competence in home language (Lindholm-Leary & Borsato, 2006; Lindholm-Leary & Genesee, 2010). Furthermore, English learners with more advanced levels of bilingual competence in both a home language and English attain significantly higher levels of academic achievement than do English learners with lower levels of bilingual competence (Lindholm-Leary & Borsato, 2006; Lindholm-Leary & Genesee, 2010). First and second language relationships have been found for language-related skills (such as depth and breadth of vocabulary), literacy-related skills (such as knowledge of the alphabet and phonological awareness), and language-processing strategies (such as inferring the meaning of new words or the use of reading-comprehension strategies).

Jensen's (2006) recent analysis of the Early Childhood Longitudinal Study-Kindergarten Cohort (ECLS-K) found that Spanish use in the classroom has a significant yet small effect (r=.2) on the mathematics achievement of SSK during the fall semester of kindergarten, such that Spanish use was associated with higher scores. Spanish-speaking kindergarteners whose teachers reported utilizing Spanish language for classroom instruction and those whose home language was Spanish scored slightly higher in mathematics (i.e. standardized tasks associated with the identification of numbers and shapes, counting, size relativity, and ordinal patterns) than Spanish-speaking kindergarteners whose teachers used only English in the classroom. This finding aligns well with the psycholinguistics literature on early cognitive and language development of language minority children of the importance of first-language maturation to underlying cognitive and second-language development during the early formative years (August & Shanahan, 2006; Jensen, 2007).

Hispanic English Learner Achievement in Mathematics and Science

The U.S Department of Education and the National Center for Education Statistics in their preliminary findings of the Early Childhood Longitudinal Study found Asian and White students had higher reading and math scores than students of other race/ethnicities including American Indian/Alaska Native, Black, Hispanic, and Native Hawaiian/Pacific Islander (Mulligan, Hastedt, & McCarroll, 2012; Mulligan, McCarroll, Flanagan, & Potter, 2014;, Mulligan, McCarroll, Flanagan, & Potter, 2015; Mulligan, McCarroll, Flanagan, & Potter, 2016; Tourangeau, Nord, Le, Wallner-Allen, Vaden-Kiernan, Blaker, 2017). Included within the aforementioned findings, students with a primary home language of English scored higher in reading and math than those coming from homes with a primary language other than English (Mulligan, Hastedt, & McCarroll, 2012; Mulligan, McCarroll, Flanagan, & Potter, 2014; Mulligan, McCarroll, Flanagan, & Potter, 2015; Mulligan, McCarroll, Flanagan, & Potter, 2016; Tourangeau, et. al 2017). Tables 1-3 highlight the underachievement of Hispanic students in the Early Childhood Longitudinal Study Cohort 2011 in the domains of Reading, Mathematics and Science. Tables 1-3 also depict those students participating in the Early Childhood Longitudinal Study whose primary language is not English. The assessment began when the children were in Kindergarten and will finalize when the children will be in their fifth-grade year. Tables 1-3 highlight the low achievement of those participating in the ECLS-K: 2011 whose primary language of the home is not English and/or who speak multiple languages in the home with no primary language identified. Tables 1-3 below summarize the scale scores among participants of the ECLS-K: 2011 in reading, mathematics and science achievement from kindergarten through third grade. Table 1-3 summarizes data by child's race and ethnicity as well as primary home language.

Table 1
Mean Scale Scores 2010-2011

	Reading		Mathematics	
	Fall 2010	Spring 2011	Fall 2010	Spring 2011
Child's race and ethnicity				
American Indian or Alaska Native, non-Hispanic	31.1	46.0	26.3	40.2
Asian	40.5	54.0	34.5	46.0
Black, non-Hispanic	32.9	47.1	25.8	37.5
Hispanic	30.3	45.3	24.7	37.8
Native Hawaiian or other Pacific Islander, non-Hispanic	32.0	48.5	27.9	41.2
Two or more races, non-Hispanic	36.1	51.0	30.5	43.2
White non-Hispanic	36.6	51.6	31.7	44.6
Primary home language				
Not English	29.4	44.2	24.1	37.3
English	35.6	50.5	30.2	42.9
Multiple home languages, no primary language identified	31.3	46.8	25.8	38.3

Table 2
Mean Scale Scores 2011-2012

	Science		Mathematics	
	Fall 2011	Spring 2012	Fall 2011	Spring 2012
Child's race and ethnicity				
American Indian or Alaska Native, non-Hispanic	24.7	27.2	48.1	62.0
Asian	23.4	26.9	56.0	68.1
Black, non-Hispanic	21.0	23.9	45.9	57.8
Hispanic	20.5	23.9	47.4	58.7
Native Hawaiian or other Pacific Islander, non-Hispanic		24.8		62.4
Two or more races, non-Hispanic	25.6	28.8	52.2	65.7
White non-Hispanic	26.0	29.2	54.4	67.5
Primary home language				
Not English	52.0	64.6	46.3	58.0
English	57.6	71.9	52.4	65.1
Multiple home languages, no primary language identified	56.1	66.0	47.4	58.7

U.S. Department of Education, National Center for Education Statistics, ECLS-K: 2011 (2014).

Table 3
Mean Scale Scores 2012-13

	Science		Mathematics	
	Fall 2012	Spring 2013	Fall 2012	Spring 2013
Child's race and ethnicity				
American Indian or Alaska Native, non-Hispanic	38.8	42.6	69.0	78.6
Asian	39.8	45.6	78.4	87.8
Black, non-Hispanic	35.9	40.3	65.2	73.6
Hispanic	35.0	40.0	67.3	77.3
Native Hawaiian or other Pacific Islander, non-Hispanic				
Two or more races, non-Hispanic	42.2	46.1	74.8	83.5
White non-Hispanic	42.7	46.6	76.6	85.3
Primary home language				
Not English	33.2	39.1	67.2	77.6
English	41.0	45.0	73.9	82.6
Multiple home languages, no primary language identified	36.1	41.3	68.5	81.1

U.S. Department of Education, National Center for Education Statistics, ECLS-K:2011 (2015).

Given the linguistic and cognitive demands non-native English speakers face, it is difficult for this particular demographic to pursue a STEM field requiring high linguistic competency and cognitive knowledge (Dang, 2015). The limited research on non-native speakers (Kanno & Harklau, 2012) has allowed for the conclusion that non-native speakers have lagged behind their English proficient peers in all content areas, specifically academic subjects that require a high demand of the English language (Abedi & Gándara, 2006), especially mathematics and science where students must be able to understand complex problems before attempting to solve them (Abedi & Lord, 2001).

Particularly in mathematics, the U.S Department of Education, Institute of Education Sciences, National Center for Education Statistics, and the National Assessment of Educational Progress (NAEP) tracked achievement level for public schools across the nation for grades 4, 8, and 12. In 2013, ELs scored significantly lower on the math assessments compared to non-ELs, where the gap was 25 points for 4th graders, 41 points for 8th graders, and 46 points for 12th graders. Similarly in 2009 EL's scored significantly lower on the science assessments compared to non-EL's, where the gap was 39 points for 4th graders, 48 points for 8th graders, and 47 points for 12th graders (NCES, 2014). The literature suggests that this performance gap could be explained by various individual and structural factors such as gender, race/ethnicity (Else-Quest et al., 2013), socioeconomic status (Krashen & Brown, 2005), immigrant generation status (Drake 2014; Rodriguez & Cruz, 2009; Rumbaut, 2005), and a host of inequitable schooling conditions (Gandara, Rumberger, Maxwell-Jolly, & Callahan, 2003)

Improved Outcomes for Students Proficient in Two Languages

For those students who have proficiency in two languages, several national and international studies have found improved mathematics achievement for bilingual students (Clarkson, 1992, 2007; Cobb, Vega, & Kronauge, 2006; Collier & Thomas, 2004; Genesse, 1983; Gomez, Freeman, & Freeman, 2006; Jensen, 2007; Lindhom-Leary & Borsato, 2001; Lindholm-Leary & Howard, 2008; Marian, Shook, & Schroeder, 2013; Thomas & Collier, 2002; Tran, Martinez-Cruz, Behseta, Ellis, Contreras, 2015). Clarkson (2007) examined a group of Australian Vietnamese students’ use of English and Vietnamese languages during mathematics classes. Specifically, the study examined the relationships between utilizing both English and Vietnamese languages and students’ achievement in mathematics. Students who used both languages interchangeably to solve mathematical word problems had generally better outcomes as compared to those utilizing only English.

Although the advantages for native bilingual speakers have been well-established in the literature (Bialystok, 1988, 1999; Luk, DeSa, & Bialystok, 2011) the potential to utilize bilingual education as a tool for enhancing cognitive development has only begun to emerge.

Executive Function Introduction

Numerous longitudinal studies suggest that executive function fosters the acquisition of emerging math skills (Blair & Razza, 2007; Bull, Espy, & Weibe, 2008; Clark, Pritchard, & Woodward, 2010; McClelland et al., 2007; Passolunghi &

Lanfranchi, 2012; Rothlisberger, Neuenschwander, Cimeli, & Roebers, 2013; Welsh, Nix, Blair, Bierman, & Nelson, 2010).

> "Executive functions perhaps make possible may of the goals we live for and permit ways to identify and achieve those goals. However, to know where one is going, it is necessary to know where you have been and where you are. In this sense, development and elaboration of executive functions are critically dependent on memory and elaboration of and, when built upon this foundation, can provide a basis for continuing adaptation, adjustment, and achievement throughout the life span" (Eslinger, 1996, p. 392).

Executive functions are concerned with the regulatory processes that allow for the initiation, modulation, and inhibition of ongoing mental attention necessary for task performance (Dempster, 1992; Dennis, 1991). Executive function skills make it possible to sustain attention, keep goals and information in mind, refrain from responding immediately, hinder distraction, tolerate frustration, acknowledge behavior consequences, and plan for the future (Zelaso, Blair, & Willoughby, 2017). Executive function regulates a person's goal-directed behavior. It contextualizes intended actions in light of past knowledge and experience, current situational cues, expectations of the future, and personally relevant values and purposes. It provides a sense of readiness, agency, flexibility, and coherence. Researchers have conceptualized executive function in terms of metacognition, inhibiting habitual responses, delay of gratification, adjusting to changing rules, and making decisions under uncertain conditions (Zelazo, Carter, Reznick, & Frye, 1997). Another definition refers to executive function as a cognitive process involved in controlling behavior and readying the person for situations. More important in real-life decision making and everyday reasoning than in responding to standardized tests, executive function comprises the ability to be mentally and behaviorally flexible to changing conditions and to provide coherence and smoothness in one's responses (Zelaso, et al, 1997).

There is a general agreement that executive function is an umbrella term for the complex cognitive processes that serve ongoing, goal-directed behaviors. In this regard, most of the definitions of executive functions include many, but not all, of the following elements: Goal setting and planning, organization of behaviors over time, flexibility, attention and memory systems that guide these processes (e.g., working memory), and self-regulatory processes such as self-monitoring. There are at least three specific EF processes: inhibitory control (IC), cognitive flexibility or shifting (CS), and working memory (WM) (Blair & Diamond, 2008; Carlson, Zelazo, & Faja, 2013; Diamond, 2013; Diamond, Carlson, & Beck, 2005; Espy, 2004; Garon, Bryson, & Smith, 2008; Harvey & Miller, 2017; Hughes, 2011; Jacques & Marcovitch, 2010; Marcovitch & Zelazo, 2009; Meuwissen & Zelazo, 2014; Senn, Espy, & Kaufmann, 2004; Wiebe et al., 2011; Zelazo, Blair, Willoughby, 2017).

This study specifically investigates two variables of EF-working memory (WM) and cognitive shifting also known as cognitive flexibility. Inhibitory control (IC) is the ability to suppress a dominant or automatic response such ignoring a distraction, stopping an impulsive utterance, or overcoming a learned response (Barkley, 1997; Clements et.al., 2015; Zelazo, Blair, & Willoughby, 2017). Cognitive flexibility refers to the ability to alternate attention or response strategies as circumstances demand-for example, considering someone else's perspective on a situation or solving a mathematics problem or solving a puzzle in multiple ways (Harvey & Miller, 2017; Zelazo, Blair, & Willougby, 2017). Working memory (WM) involves keeping information in mind and utilizing and manipulating it for the purpose of carrying out a task (Harvey & Miller, 2017; Zelazo, Blair, & Willougby, 2017). Working memory may assist a child during the acquisition of number facts and application to multistep problem solving in mathematics. This study specifically focuses on two variables of EF-working memory (WM) and cognitive shifting also known as cognitive flexibility.

Figure 11. *Semantic Map of Executive Function and Related Terms*. Note: Near synonyms of EF include: cognitive control, executive attention, executive control, executive functioning, and fluid abilities).

Executive Functioning and School Success

EF processes have been cited as "crucial building blocks for the early development of both cognitive and social capacities" (p. 3) in a joint paper by the Policy and Procedures (Center on the Developing Child at Harvard University, 2011). In the past two decades, EF skills have emerged as a major focus of research in the fields of psychology, neuroscience, and education (Zelaso, Blair, & Willoughby, 2017). Executive function skills help individuals focus on and persist in the attainment of goals, which are both critical components to academic success (Little, 2017). Executive function is strongly associated with school success (Agostino, Johnson, & Pascual-Leone, 2010; Blair, Razza, 2007; Diamond, Barnett, Thomas, & Munro, 2007; Dorney, 2005; Van der Ven, Kroesbergen, Boom, & Leseman, 2012).

Measures of executive function have been found to be better indicators of later academic achievement as compared to IQ or entry-level reading and math skills (Bull & Scerif, 2001; Esposito & Baker-Ward, 2013; McClelland & Morrison, 2000; Waber, Gerber, Turcios, Wagner, & Forbes, 2006). Research has established that EF skills provide a foundation for learning in school settings and is central to school readiness and early school achievement (Blair 2002; Blair and Raver, 2015).

Executive function deficits have been identified as potential targets of early intervention efforts designed to help young children with either mathematics or reading difficulties (Bull, Espy, & Wiebe, 2008; Bull & Scerif, 2001; Pham & Hasson, 2014: Pickering & Gathercole, 2004; Swanson & Saez, 2003; Swanson, Zheng, & Jerman, 2009; Toll, van der Ven, Kroesbergen, van Luit, 2011; Van der Ven, Kroesbergen, Boom,

& Leseman, 2012). Deficits in executive functions have been hypothesized to interfere with young children's success in school, including understanding instructions as well as managing and selectively ignoring simultaneous stimuli on their attention while completing assigned work (Alloway, Gathercole, Kirkwood, & Elliot, 2009; Gathercole, Durling, Evans, Jeffcock, & Stone, 2008). According to Morgan, Li, Farkas, Cook, Pun, & Hillemeier (2017), "direct observation finds that children with executive functioning deficits more frequently fail to complete multi-step instructions by their teachers and to finish complex tasks. Deficits in executive functions are also thoughts to interfere with children's mathematics as well as reading achievement (Swanson & Beebe-Frankenberger, 2004).

Improved Executive Functioning in Bilingual Students

Research conducted during the past two decades has found advanced levels of bilingual competence are linked with several significant cognitive advantages (Bialystok 2001, 2007, 2008). Ardila (2006) proposed that the development of language is the primary contributor of cognitive growth. The ability of a language to provide mental representations to our perception and knowledge of the world allows better understanding of the environment, enabling our survival. Knowledge of brain development and neuronal connections suggests that there is an optimal period of time for language development, which has been found regardless of the language being learned (Vega, 2008). Even when a second language is learned after the critical window of six or seven years of age, research has shown that the development of neuronal systems sub-serving the second language will enhance executive functions (Ratey, 2002).

In relating use of another language and executive function, Clarkson (2007) asserts that the evidence that bilingual young people, relative to monolingual controls, show greater cognitive flexibility, creativity, divergent thought (Dorney, 2005), and improved problem solving abilities, is very significant.

Children having the capacity to master two languages have two or more words for each object and idea, associate diverse meanings to words, and can therefore develop the ability to think more flexibly about the world surrounding them. Bilingual students possess certain cognitive advantages related to executive function, also known as executive control. Some of advantages described by Lindholm-Leary and Genesee (2010) include; completion of tasks, problem solving, cognitive abilities related to attention, inhibition, monitoring, and switching focus of attention. Collectively, these cognitive skills are referred to as executive function (executive control) and are located in the frontal lobe regions of the brain (Lindholm-Leary & Genesee, 2010).

Bialystok's (2001) book *Bilingualism in Development: Language, Literacy and Cognition* discussed the effects of bilingualism on non-linguistic aspects of children's cognition, such as quantity, number, problem solving, and sorting, all mathematical skills linked to achievement. Bialystok noted the advantage of bilingualism in cognitive tasks requiring high levels of inhibition or conceptual demands including misleading information needing to be filtered.

Carlson and Meltzoff's (2008) study of kindergarteners using nine different measures of executive functioning, found comparable scores in executive functioning tasks between Spanish-English bilinguals and English monolinguals, despite lower socioeconomic status among the bilingual group. In this particular study, even when parental education, children verbal skills, and child age variables were controlled, lower socioeconomic status bilingual children outperformed their middle SES English monolingual counterparts in inhibitory control and working memory executive function tasks.

Recently, Bialystok & Barac (2012) provided evidence on the benefits of second-language exposure on EF may not be limited to native bilinguals but has the possibility of being acquired through an immersion education model. Their study, consisting mostly of upper middle class English speaking children in minority-language instruction, concluded that the number of years elementary school children spent in Hebrew and French immersion education predicted their performance on measures of EF.

Esposito & Baker-Ward (2013) examined the benefits of dual Spanish-English (50:50) immersion model on EF in low-income elementary school children. Their sample consisted of 120 ethnically diverse, low-income children from grades K, 2^{nd}, and 4^{th} enrolled in both dual language and traditional classrooms. The researchers compared the performance of children in traditional classrooms in which instruction was delivered only in English with that of children in a 50:50 dual language education model (Esposito & Baker-Ward, 2013). Their results suggest that both native English and Spanish-speaking dual-language students experience the same benefits as do minority-language immersion students (those educated through a minority language) for aspects of EF. Dual-language education offers minority-language students a facilitated approach to quickly acquiring the majority language; provides the benefits of learning two languages to both majority and minority language students; and may enhance the ability for both students to regulate attention (EF).

Executive Function and Achievement

Numerous longitudinal studies suggest that executive function fosters the acquisition of emerging math skills (Blair & Razza, 2007; Bull, Espy, & Weibe, 2008; Clark, Pritchard, & Woodward, 2010; McClelland et al., 2007; Passolunghi & Lanfranchi, 2012; Rothlisberger, Neuenschwander, Cimeli, & Roebers, 2013; Welsh, Nix, Blair, Bierman, & Nelson, 2010).

Little's (2017) analysis of the ECLS-K: 2011 found that Black and Hispanic students entering kindergarten had significantly lower working memory and cognitive flexibility skills than White students. On the Numbers Reversed task, Hispanic students enter kindergarten performing 0.59 standard deviations lower, on average, than their White peers. In the same study, the author focused on socioeconomic status, the analysis reveals large gaps in working memory and cognitive flexibility at school entry (Little, 2017). Students in the top socioeconomic status quintile score 1.01 standard deviations higher, on average, compared to their peers from the lowest socioeconomic status quintile in the Numbers Reversed task. Little's (2017) method focuses on modeling standardized outcome measures using ordinary least squares (OLS) regression. Each dependent variable (Numbers Reversed, DCCS, math, and reading) was regressed on indicators of race and socioeconomic status in the fall of kindergarten and the spring of kindergarten, first and second grades. Models for each outcome and assessment wave were estimated including indicators of race and socioeconomic status together (Little, 2017)

Morgan et al. (2016) analyzed a nationally representative and longitudinal sample of the ECLS-K: 2011 and found that executive functioning deficits are uniquely predictive of kindergarten children's risk for later experiencing learning difficulties. Their results indicate that kindergarten children with working memory and cognitive flexibility deficits were at increased risk of experiencing reading and mathematics difficulties in first grade. Their findings provide empirical support for experimental evaluations of school-based, multi-component interventions designed to address early onset of learning difficulties through the remediation of executive function deficits. Their study found that deficits in working memory are more strongly predictive of experiencing learning difficulties in childhood and that the relation between executive function deficits is relatively stronger for mathematics than for reading achievement.

Morgan et al. (2017) performed multivariate logistic regression of the ECLS-K: 2011 using a multi-year panel design, multiple criterion and predictor variable measures, extensive statistical control for potential confounds including autoregressive prior histories of both reading and mathematics difficulties, and additional epidemiological methods to preliminarily examine hypothesized relations.

Link Between Executive Function, Language, and Math Achievement

Over the past decade, research associating executive functioning with early academic achievement in mathematics has begun to emerge (Allan, Hume, Allan, Farrington, & Lonigan, 2014; Blair and Diamond, 2008; Blair & Razza, 2007; Espy et al., 2004; Lonigan, Allan, Goodrich, Farrington, & Phillips, 2015; McClelland, Cameron, Connor et al., 2007; McClelland, Cameron, Wanless, et al., 2007).

Bilingual children exhibit increased non-linguistic, cognitive benefits of executive functioning skills (Bialystok, 2007; Bialystok & Martin, 2004; Bialystok & Viswanathan, 2009; Carlson & Meltzoff, 2008), which correlate with math performance (Blair & Razza, 2007; Bull, Espy, Wiebe, 2008; Marian, Shook, Schroeder, 2013; Mazzoco & Kover, 2007; McClelland et al., 2007; Passolunghi & Siegel, 2001). Researchers Bull and Scerif (2001) assessed third graders with executive functioning tasks (e.g. card sorting task) and with a mathematics test of addition and subtraction; and through multiple linear regression analyses, they found that executive functioning reliably predicted mathematics performance. Executive function promotes performance on mathematics problems by "allowing the student to hold the problem in working memory, to shift one's focus between different aspects of the problem and different approaches to the problem, and to suppress a tendency to respond to salient but irrelevant elements of the problem" (Marian, Shook, & Schroeder, 2013, p. 180). Additionally, executive functioning enhances students' focus on mathematics and science lessons during class instruction, which may result in enhanced acquisition of math and science skills.

Blair and Razza's (2007) study with preschool aged children in Head Start found that inhibitory control skills (one measure of executive function) was a significant predictor of phonemic awareness, letter knowledge, and mathematics skills. Their study found a higher effect size in mathematical skills than for literacy-related skills. Similar results were found in a recent meta-analysis of 75 peer-reviewed studies of preschool and kindergarten children from households of varying SES (Allan et al, 2014). The aforementioned research found that Inhibitory Control skills were moderately associated with academic skills, with a higher effect size for math (*r*=.34) than for literacy outcomes (*r*=.25). The positive links between executive functioning skills and achievement in math may help to explain academic achievement gaps across children with different language experiences. Given the aforementioned literature findings, the overarching goal of the present study is to use theoretically driven covariates- Spanish of the home and executive functioning to explain the variation in initial levels (intercepts) and slopes (over four time points) across students.

For school leaders facilitating systemic change and overall school improvement, improved acquisition of mathematics and science skills improves performance in state and federal standardized assessments. This study attempts to (change wording) relate the outcomes of language and executive functioning in working memory and cognitive flexibility as they relate to mathematics and science achievement.

Summary of Literature Reviewed

Chapter 2 in this proposed study described more in detail the condition of Hispanics/Latinos in the United States and their classification as linguistic minorities. Followed by the literature on home language and academic achievement, specifically Spanish was also evaluated, and the connection was made to STEM outcomes for Hispanic linguistic minorities. The link between proficiency in two languages and improved outcomes was examined. In addition, the definition of executive functioning and its roles in school success, and specifically for bilingual students was examined. Finally, the link between executive function, language and achievement was analyzed in the literature.

III. METHODS

Introduction

Scores on tests of educational achievement may be influenced by two or more co-occurring variables (a.k.a. covariates). In this study, co-occurring variables related to educational achievement in mathematics and science include Spanish spoken at home (i.e. English as a second language rather than primary) and assessments of executive function. The existence of co-occurring factors and their relationship to how students' mathematics and science achievement scores change over time presents a substantial challenge to providing effective student-centered interventions. The existence of co-occurring factors and their relationship to how students' mathematics and science achievement scores change over time presents a substantial challenge to providing effective student-centered interventions. To increase understanding about the complex relations between student-level covariates and achievement scores, we introduce a parallel process latent growth model (PPLGM) with covariates.

Traditionally, repeated measures analysis of variance (ANOVA) has been used to examine changes in individuals across time in experimental and quasi-experimental designs. Hypothesis tests can be conducted to determine whether the means of the variables measured at different time points are equal or to determine the shape of the growth, or change, trend (linear, quadratics, exponential). Researchers have identified several limitations associated with the repeated measures ANOVA (Hox, 2010; Raudenbush & Bryk, 2002; Raykov & Marcoulides, 2006). These shortcomings include the inability to examine random variation in growth rates, in measurement error, problematic assumptions such as the normality of the data between-subject factors, and the equality of the covariance matrices across measurements.

Cross-sectional studies focus on outcome and process variables at one point in time and are not well suited for investigations of processes that are assumed to be dynamic (Heck & Thomas, 2015). Therefore, it is difficult to establish proper time-ordering necessary in addressing causal relationships in cross-sectional analyses. A major benefit of analyzing longitudinal data is the ability to disentangle causal relationships while examining change processes at the student level. Growth curve analyses are one of several techniques used to address the goals of longitudinal research, the goals being to analyze both intra-individual change and inter-individual differences in intra-individual change (Zhang, McArdle & Nesselroade, 2012). Longitudinal growth curve analysis allows for the measurement of both group and individual variation in growth (Fan & Konold, 2009) through estimation of 1) linear and non-linear slopes that model the rate of change across time, 2) mean intercept and slope values reflecting the group average at initial status (Time 1) and the average rate of growth across individuals, respectively, 3) individual variation in intercepts and slopes, and 4) the correlation between the intercept and slope.

Structural Equation Modeling Perspective

From the SEM perspective, conventional growth curve modeling is specified in two parts: 1) measurement part that links repeated measures of an outcome to latent growth factors 2) a structural part that links latent growth factors to each other and to individual-level predictors (Kaplan, 2002). If the outcome variable is defined as a p –

dimensional vector **y**, following Muthén (2002) and Kaplan (2002), the measurement part of the model can be expressed as

$$\mathbf{y}_i = \mathbf{v} + \Lambda\eta_i + \mathbf{K}\mathbf{x_i} + \mathbf{e}_i$$

Where, the *p*-dimensional vector $\mathbf{y}_i$ representing the empirical growth for child *i*; **v** is a *p*-dimensional parameter vector of measurement intercept; Λ is a *p* x *m* matrix of factor loadings; η *p*-dimensional parameter vector of latent variables; **K** is a *p* x *q* parameter matrix of regression slopes; **x** is a *q*-dimensional vector of covariates; and **e** is a *p*-dimensional vector of residuals. The structural part of the model is defined in terms of the latent variables regressed on each other and the *q*-dimensional vector **x** of independent variables,

$$\eta_i = \boldsymbol{a} + \boldsymbol{B}\eta_i + \mathbf{\Gamma}\mathbf{x_i} + \xi_i$$

Where η is defined as before; $\boldsymbol{a}$ is a *m*-dimensional vector that contains the population initial status and growth parameters $\mu_{\pi 0}$ and $\mu_{\pi 1}$; **B** is an *m* x *m* matrix containing regression slopes that relate the latent variables to each other; **Γ** is an *m* x *q* matrix of regression coefficients relating the latent growth factors to the independent variables; and ξ is an *m*-dimensional vector of residuals. The basic latent growth curve models formulated as a single-level (multivariate) model can also be extended to include situations where we wish to include group-level variables in modeling change.

Latent Growth Models

The term latent growth model (LGM) or latent growth curve (LGC) modeling within the framework of structural equation modeling (SEM) is now considered one of the most powerful and informative approaches to the analysis longitudinal data (Curran & Hussong, 2003). This method enables researchers to test for differences in developmental trajectories across time, conventional repeated measures analyses such as ANOVA, ANCOVA, and MANOVA fail to provide this opportunity (Byrne, 2010). Although the aforementioned strategies are capable of describing an individual's developmental trajectory, they are incapable of capturing individual differences in these trajectories over time (Curran & Hussong, 2003; Duncan & Duncan, 1995; Fan, 2003; Willet & Sayer, 1994).

Outlined below are the requirements for the analysis of an LGM in SEM (Duncan, Duncan, Strycker, Li, & Alpert, 1999).

1. A continuous dependent variable measured on at least three different occasions.
2. Scores that have the same units across time and can be said to measure the same construct at each assessment.
3. Data that are time structured, which means that cases are all tested at the same intervals and those intervals do not have to be equal.

As noted by Bauer (2003), Curran (2003), and others, latent growth models analyzed in SEM are in fact multilevel (two-level) models that explicitly acknowledge the fact that scores are clustered under individuals (repeated measures). Scores from the same case are probably not independent and this lack of independence must be considered in the statistical analysis. Latent growth models are often analyzed in two steps. The first concerns a change model of just the repeated measures variables. This model attempts to explain the covariances and means of these variables. Given an acceptable change model, the second step adds variables to the model that may predict change over time. This two-step approach makes it easier to identify potential sources of poor model fit compared with the analysis of a prediction model in a single step.

Current Study

Examination of the research questions above implies a need for a growth curve model that adequately accounts for change over time in mathematics, and science. Whether and to what extent executive functioning deficits are uniquely associated with kindergarten, first grade, second grade, and third grade risk of later experiencing science and mathematics difficulties. A population-based and multi-year longitudinal sample of U.S. schoolchildren, including two measures of two types of executive function (working memory, cognitive flexibility).

Conducting longitudinal growth curve modeling involves 2 stages. At stage 1, the goal is to study individual student change over time. Each student has a regression line or a growth curve that plots the variable of interest over time see Figure 12.

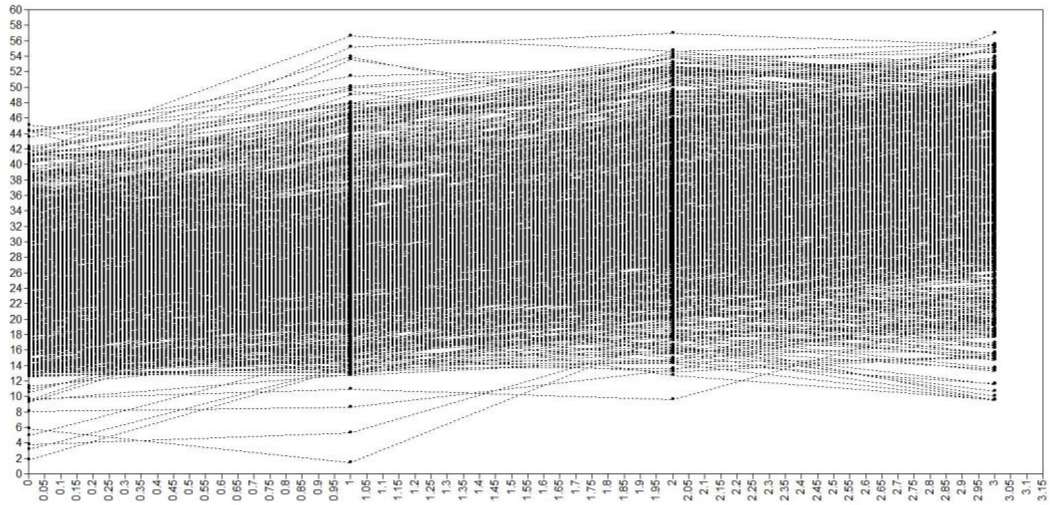

Figure 12. *Change Over Time In 1000 Randomly Selected Students ECLS-K: 2011 Data*

At Stage 2, the individual level regression equations can then be summarized to obtain an average intercept (or initial level) and average slope, or average rate of change for all students. The intercept and slope each have their own variance. Interpreting the variance in the intercept and slope provides a way to evaluate the magnitude of variability in student scores at each time point (i.e. intercept variance) and the degree of change between time points (i.e. slope variance over time).

The overarching goal here is to use theoretically driven covariates. For example, in this study, Spanish spoken in the home and executive functioning explain the variation in initial levels (intercepts) and slopes (over four time points) between students. In this investigation, student achievement trajectories in mathematics and science will be modeled across four time points. Scores used occurred at one semester intervals beginning at 4.5 years of age or kindergarten (Time 1), Spring Kindergarten (Time 2) Fall of First Grade (Time 3) and Spring of First Grade (Time 4). Our PPLGM (Figure 1) utilizes longitudinal growth modeling based on three types of analyses each serving a different purpose. The first univariate analysis models growth of mathematics or science in isolation. In the second analysis, an unconditional multivariate analysis, examines the joint associations between growth factors of mathematics and science. Finally, a conditional multivariate analysis incorporates time-varying covariates as predictors of mathematics achievement scores at each year. In our analyses, we allow measures assessed within the same occasion to covary (Blozis, Harring, & Mels, 2008).

In Figure 1, our PPLGM allows for estimation and specification of both latent intercept and latent slope terms, as illustrated by the ellipses. The specification of a latent intercept provides for measurement of initial status (i.e. achievement status at 54 months of age), and the latent slope yields a measure of growth across time. All models were estimated with M*p*lus 7.0 program using full information maximum likelihood estimation. Four measures of fit were considered in evaluating model quality. Four measures of fit were considered in evaluating model quality. These include the Bentler-Bonnett normed fit index (NFI), Tucker-Lewis index (TLI), comparative fit index (CFI), and root mean square error of approximation (RMSEA). Research suggests that better fitting models produce values around .95 (Hu & Bentler, 1999). Alternatively, smaller RMSEA values support better fitting values of .08 or less indicating good fit (Fan, Thompson, & Wang, 1999).

Instrumentation

The Early Childhood Longitudinal Study is currently sponsored by the National Center for Education Statistics (NCES), within the U.S. Department of Education's

Institute of Education Sciences, to provide detailed information on the school achievement and experiences of students throughout their elementary school years (McCarroll, Flanagan, & Potter, 2016). The ECLS-K: 2011 is the third iteration in a series of longitudinal studies of young children (ECLS-B and ECLS-K in 1998). The ECLS-K:2011 was created in order to advance research possibilities by providing updated information and addressing recent changes in education policy that were not measured fully in the previous studies. Significant changes since the first initial ECLS-K in 1998 include the passage of No Child Left Behind (NCLB) legislation, a rise in school choice, and an increase in English Language Learners. The ECLS-K: 2011 is utilized in this study due to its rich, reliable data and its longitudinal nature that allows to examine a wide range of school and cognitive variables. The students participating in the ECLS-K: 2011 were assessed longitudinally from kindergarten (the 2010-2011 school year) through the spring of 2016, when most are expected to be in fifth grade. The ECLS-K: 2011 places an emphasis on measuring students' experiences within multiple contexts and development in multiple domains including cognitive, socio-emotional, and physical development through direct and indirect methods. The ECLS-K: 2011 is the first nationally representative and longitudinal study to assess children's executive function skills. The ECLS-K: 2011 kindergarten direct cognitive assessment battery was designed to assess kindergartners' knowledge and skills in reading, mathematics, and science (Mulligan, Hastedt, & McCarroll, 2017). Because of the ECLS-K:2011 is a longitudinal study, the assessments also were designed to allow for the measurement of growth in these domains across time. The longitudinal design of the ECLS-K:2011 required that the cognitive assessments be developed to support the measurement of change in knowledge and skills demonstrated by demonstrated by children from kindergarten entry through the spring of fifth grade. The ECLS-K:2011 reading and math specifications were based on the frameworks developed for the National Assessment for Educational Progress. Although the NAEP assessments are administered starting in fourth grade, the specifications were extrapolated down to kindergarten based on current curriculum standards from several states and, for math, the National Council of Teachers of Mathematics *Principles and Standards for School Mathematics.* The frameworks necessarily cover content strands applicable to a range of content at different grade levels, for example from number sense (i.e., basic knowledge of numbers) to algebra in mathematics. Content appropriate for most kindergartners was included in the kindergarten assessments. For example, in the kindergarten math assessment, the "algebra" content strand was assessed through children's recognition of patterns. While the assessments were designed to contain mostly items that assessed knowledge and skills at a kindergarten level, easier and more difficult items were included to measure the abilities of students performing below or above grade level.

A nationally representative sample of approximately 18,080 children from about 1,310 schools participated in the base-year administration of the ECLS-K: 2011 in the 2010-11 school year. The direct child assessments are cognitive assessment batteries and socioemotional items developed specifically for use in the ECLS-K: 2011 and administered directly to the children (Tourangeau, Nord, Lê, Wallner-Allen, Vaden-Kiernan, Blaker, & Najarian, 2017). The direct cognitive assessments were designed to measure children's knowledge and skills at given time points, as well as track their academic growth in different subject areas across time. Results from the assessments for reading, mathematics, and executive function (working memory and cognitive flexibility) enable researchers to measure growth from the fall of children's kindergarten year (fall 2010) through the spring of 2016, when most ECLS-K:2011 students were in the fifth grade (Tourangeau et al., 2017). Science knowledge and skills can be examined beginning in the spring of the children's kindergarten year. The ECLS-K: 2011 assessed knowledge and skills that are typically taught and developmentally important. The assessment frameworks were derived from national and state standards, including those of the National Assessment of Educational Progress (NAEP), the ECLS-K frameworks, and selected states' curriculum standards (Tourangeau et al., 2017).

Table 4

Instruments Used for Data Collection: 2010-2013

Child Assessment Instrument	Fall Kindergarten	Spring Kindergarten	Fall First Grade	Spring First Grade	Fall Second Grade	Spring Second Grade
Mathematics	X	X	X	X	X	X
Executive Function	X	X	X	X	X	X
Science		X	X	X	X	X

U.S. Department of Education, National Center for Education Statistics, ECLS-K:2011 (2017).

Sampling Procedure

The ECLS-K:2011 cohort was sampled using a multistage sampling design to produce national-level estimates where there is approximately the same probability for each child to be selected. In the first stage, 90 primary sampling units (PSUs) were selected from a national sample of PSUs, or geographic areas based on the 2007 Census Bureau population estimates (Mulligan, Hastedt, & McCarroll, 2012; Mulligan, McCarroll, Flanagan, & Potter, 2014; Mulligan, McCarroll, Flanagan, & Potter, 2015; Mulligan, McCarroll, Flanagan, & Potter, 2016; Tourangeau, et. al 2017). The PSUs were counties and county groups from the 3,141 total counties in the United States. In the second stage, public and private schools educating kindergartners (or ungraded schools educating children of kindergarten age) were selected within the PSUs. Primary sampling units and schools were selected within the sampled PSU's and with probability proportional to the population size. This included oversampling of Asians, Native Hawaiians and other Pacific Islanders. The schools were selected from a preliminary version of the frame developed for the 2010 National Assessment of Educational Progress (NAEP), which contained information about public schools that were included in the 2006-07 Common Core of Data and private schools were selected using the Private School Survey (2008-08 PSS). In the third stage of sampling, approximately 23 kindergarteners were selected from a list of all enrolled kindergarteners (or students of kindergarten age being educated in an ungraded classroom) in each of the sampled schools.

Sampling Errors and Weighting

The ECLS-K:2011 data are weighted to compensate for unequal probabilities of selection at each sampling stage and to adjust for the effects of school, teacher, before and after school care provider, child, and parent non-response (Mulligan, Hastedt, & McCarroll, 2012). The sample weights to be used in ECLS-K:2011 analyses were developed in several stages. The first stage of the weighting process assigned weights to the sampled primary sampling units that are equal to the inverse of the PSU probability of selection (Mulligan, et. al, 2017). The second stage of the weighting process assigned weights to the schools sampled within selected PSUs. The base weight for each sampled school is the PSU weight multiplied by the inverse of the probability of selecting the school from the PSU. The base weights of responding schools were adjusted to compensate for nonresponse among the set of eligible schools. These adjustments were made separately for public and private schools.

The Sample

The Early Childhood Longitudinal Study, Kindergarten Class of 2010-11 (ECLS-K:2011) is providing national data on children’s characteristics as they progress from kindergarten through the 2015-2016 school year, when most of the children are in fifth grade. From the publicly available dataset (National Center for Education Statistics, 2017), a total of 1,221 clusters of schools were originally selected for the ECLS-K:2011, of which 1,003 were clusters of public schools and 218 were clusters of private schools. For purposes of this current study, 218 clusters of private schools were filtered out. Of the original sample of 18,174 children in the original dataset, an unweighted sample of

2,272 remains as the original sample was filtered and limited to those students attending public schools. Cases in private schools were deleted. The sample was also limited to those in city and suburban areas, rural and town were deleted.

Due to unequal selection probability occurring when elements in the population are sampled at different rates (Stapleton, 2002), weights were applied to the current sample in order to compensate for the unequal probability of selection, non-response and noncoverage, and poststratification (Kalton, 1989). Populations that are often oversampled (such as American Indians/Alaska Natives) in national datasets have a smaller weight value (Thomas & Heck, 2001). Ignoring disproportionate sampling may result in biased parameter estimates and poor performance of test statistics and confidence intervals (Pfeffermann, 1993) as the weights are required to produce estimates that are representative of the intended population (U.S. Department of Education, 2002). As for the current study, replicate weights W1C0 W1CI were designed for the analysis of child direct assessment data from fall and spring kindergarten, fall first grade, alone or in combination of a limited set of child characteristics (age, sex, race-ethnicity). Once the aforementioned weights were applied, the new sample size became 629,706 students. Although our original unweighted sample of consisted of 2,272 students, once the weights were applied, the sample is now generalized to a much larger population.

Assessment Timing

An important issue to be considered when analyzing achievement scores and gains is assessment timing: children's age at assessment, the date of assessment, and the time interval between assessments. Most sampled children were born throughout the second half of 2004 and first half of 2005, but their birth dates were not related to testing dates. As a result, children were tested at different developmental and chronological ages. Assessment dates ranged from August to December for the fall data collections, and from March to June for the spring data collections. Children assessed later in a data collection period in a particular grade level, for example in December during a fall collection, may be expected to have an advantage over children assessed earlier in the data collection period, for example in the first days or weeks of school, because they had more exposure to educational content before being assessed. Substantial differences in the intervals between assessments may also affect analysis of gain scores. Children assessed in September for the fall data collection and June for the spring data collection have more time to learn knowledge skills than do children assessed first in November and then again in March. These differences in interval may or may not have a significant impact on analysis results. In designing an analysis plan, it is important to consider whether and how differences in age, assessment date, and interval may affect the results; to look at relationships between these factors and other variables of interest; and to adjust for differences, if necessary. When using the IRT scale scores as longitudinal measures of overall growth, analysts should keep in mind that gains made at different points on the scale have qualitatively different interpretations. Children who made gains toward the lower end of the scale, for example, in skills such as identifying letters and associating letters with sounds, are learning different skills than children who made gains at the higher end of the scale, for example, those who have gone from reading sentences to reading passages, although their gains in number of scale score points may be the same.

Comparison of gains in scale score points is most meaningful for groups that started with similar initial status. One way to account for children's initial status is to include a prior round assessment score as a control variable in an analytic model. For example, the fall kindergarten scale score could be included in a model using the spring kindergarten scale score as the outcome

Measures and Covariates

Executive functions are interdependent processes that work together to regulate and orchestrate cognition, emotion and behavior and aide child learning in the classroom. Two measures of executive function were included in the kindergarten, first-grade, and second-grade assessment battery: the *Dimensional Change Card Sort* (DCCS) (Zelazo, 2006; Zelazo et al., 2013), assessing children's cognitive flexibility, and the Numbers Reversed subtest of the *Woodcock-Johnson III* (WJ III) *Tests of Cognitive Abilities* (Woodcock, McGrew, and Mather 2001), assessing working memory. The same versions of the DCCS and the Numbers Reversed tasks were administered in fall and spring of first grade. In second grade, the DCCS was changed to computerized administration to remain age appropriate, while the version of the Numbers Reversed task remained the same version used in the earlier data collection rounds. Detailed information regarding administration and psychometric properties is available in the ECLS-K:2011 user's manual (Tourangeau et al., 2017). The last covariate included in the present study is whether Spanish is spoken in the home.

ECLS-K: 2011 Test Fairness

ECLS-K:2011 child assessors were trained on general issues for working with children in the particular age group covered by the study's round. There was no component in their trainings that specifically addressed cultural proficiency for either the child assessors or parent interviewers. However, it is of note that most of the field staff were former teachers so had experience working with the population. Additionally, local staff was hired to each of the primary sampling units (PSU) so field staff were working in their region, which could have contributed to greater cultural fluency in the area for which they were working. They also had bilingual assessors available for the Spanish parent interviews and, in the earlier rounds when it was administered, the Spanish child assessment components.

Differential Item Functioning

Differential item functioning was analyzed as part of the ECLS-K:2011 field tests for the child assessments. Future efforts include having a detailed discussion of the field test work and the subsequent development of the child assessments in our forthcoming psychometric reports. Unfortunately, at this there is no estimated release date for those reports, so for now the greatest detail on assessment development can be found in chapters 2 and 3 of the user's manuals (Tourangeau, et al., 2017).

Data Collection

The released public use ECLS-K: 2011 will include 6 rounds of data collection: Fall and Spring Kindergarten, Fall and Spring First, and Spring Third, and Spring Fifth. To conduct the conventional growth curve model, eight time points were chosen in this study: Fall Kindergarten, Spring Kindergarten, Fall First, Spring First, Fall Second, Spring Second, Fall Third, and Spring Third. Fall and spring data collections were conducted in the 2010-2011 school year, when the sampled children were in their kindergarten year; in the 2011-2012 school year, when most sampled children were in first grade; in the 2012-2013 school year, when most study children were in second grade; and in the spring of 2014 when most children were in third grade. The publicly available data used for purposes of this study end in 2014 when the children are in third grade.

Table 5
School Year, Grade, and Data Collections ECLS-K:2011

School Year	Grade	Data Collections
2010-2011	Kindergarten	Fall 2010, Spring 2011
2011-2012	First Grade	Fall 2011, Spring 2012
2012-2013	Second Grade	Fall 2012, Spring 2013
2012-2014	Third Grade	Spring 2014
2014-2015	Fourth Grade	Spring 2015
2015-2016	Fifth Grade	Spring 2015, Spring 2016

U.S. Department of Education, National Center for Education Statistics, ECLS-K:2011 (2017).

Direct Child Assessment

In the second-grade collections, children were assessed in reading, mathematics, and science in both the fall and the spring. The majority of the items included in the second-grade assessments had been included in the first-grade assessments. However, to ensure that the assessments adequately measured the knowledge and skills of the children as they progressed through school, new, more difficult items were added to the assessments in second grade, and easier items reflecting lower level first-grade skills were omitted. All children received the assessments designed for the second-grade collections, regardless of their actual grade level. In both the fall and the spring, students' executive function skills were assessed and their height and weight were measured. The assessments were administered directly to the sampled children on an individual basis by trained and certified child assessors. The battery of assessments was designed to be administered within about 60 minutes per child.

Research Variables

Table 6
Direct Assessment Research Variables

Variable Name	Description	Measures in this Study	Parallel Growth Curve Model Symbol
X1MSCALK2	F 2010 Math IRT Scale Score	Math Achievement	Math 1
X2MSCALK2	S 2011 Math IRT Scale Score	Math Achievement	Math 2
X3MSCALK2	F 2011 Math IRT Scale Score	Math Achievement	Math 3
X4MSCALK2	S 2012 Math IRT Scale Score	Math Achievement	Math 4
X2SSCALK2	S 2011 Science IRT Scale Score	Science Achievement	Science 1
X3SSCALK2	F 2011 Science IRT Scale Score	Science Achievement	Science 2
X4SSCALK2	S 2012 Science IRT Scale Score	Science Achievement	Science 3
X5SSCALK2	F 2012 Science IRT Scale Score	Science Achievement	Science 4
X1NRWABL	F 2010 Numbers Reversed W-Ability Score (Executive Function)	Covariate	NR 1

Table 6 Continued

X2NRWABL	S 2011 Numbers Reversed W-Ability Score (Executive Function)	Covariate	NR 2
X3NRWABL	F 2011 Numbers Reversed W-Ability Score (Executive Function)	Covariate	NR 3
X4NRWABL	S 2012 Numbers Reversed W-Ability Score (Executive Function)	Covariate	NR 4
X1DCCSTOT	F 2010 Dimensional Card Sort Combined Score (Executive Function)	Covariate	CS 1
X2DCCSTOT	S 2011 Dimensional Card Sort Combined Score	Covariate	CS 2
X3DCCSTOT	F 2011 Dimensional Card Sort Combined Score	Covariate	CS 3
X4DCCSTOT	S 2012 Dimensional Card Sort Combined Score	Covariate	CS 4
C1SPHOME	F 2010 Speak Spanish at Home	Covariate	Sp 1
C2SPHOME	S 2011 Speak Spanish at Home	Covariate	Sp 2
C3SPHOME	F 2011 Speak Spanish at Home	Covariate	Sp 3
C4SPHOME	S 2012 Speak Spanish at Home	Covariate	Sp 4

Table 6 above describes the different variables, their names in the ECLS-K: 2011 dataset, the description of each, the type of variable used in the study, the symbol in the proposed research diagram and longidinal growth curve model. Table) 7 below also describes the variables according to locale such as school type by city/suburb, when the variables were measures (i.e., Fall 2010/Fall 2011, Spring 2011, Spring 2012).

Table 7
Research Variables by Location and School Type

Variable Name	Description	Measures in this Study
X1LOCALE	Fall 2010 Location type of school	City/Suburb
X2LOCALE	Spring 2011 Location type of school	City/Suburb
X3LOCALE	Fall 2011 Location type of school	City/Suburb
X4LOCALE	Spring 2012 Location type of school	City/Suburb
X1PUBRI	Fall 2010 Public or private school	Public
X2PUBRI	Spring 2011 Public or private school	Public
X3 PUBRI	Fall 2011 Public or private school	Public
X4PUBRI	Spring 2012 Public or private school	Public
W1C0	W1 C1 Child Full Sample Weight	
W1C1-W1C80	Child Replicate Weights	

U.S. Department of Education, National Center for Education Statistics, ECLS-K:2011 (2017).

IRT Scores

Broad-based scores using the full set of items administered in the kindergarten and first-grade assessment in reading, math, and science were calculated using item response theory (IRT) procedures. IRT is a method for modeling assessment data that makes it possible to calculate an overall score for each domain measured (Mulligan, McCarroll, Flanagan, & Potter, 2014). Similar to other methods, “IRT is a model-based theory of statistical estimation that conveniently places persons and items on the same metric based on the probability of response outcomes” (Price, 2017, p. 330) having the goal of constructing models of behavior and/or performance in relation to theory. IRT's postulates that a latent trait is represented as a continuum along a measurement scale and are able to predict or explain either a social, behavioral or psychological attribute (Price, 2017). IRT originates from the pattern of examinees' responses to a set of test items. IRT scales scores in mathematics and science scores served as the outcome measures.

This method was used to calculate an overall score for the ECLS-K:2011 because the study employed a two-stage assessment in which children were administered a set of items appropriate to their demonstrated ability level rather than all of the items in the assessment. Although this procedure resulted in children being administered different sets of items, there was a subset of items that all children received (the items in the routing tests, plus a set of items common across the different second stage forms). These common items were used to calculate scores for all children on the same scale.

IRT has several advantages over raw number-right scoring. By using the overall pattern of right and wrong responses and the characteristics of each item to estimate ability, IRT can adjust for the possibility of low-ability child guessing several difficult items correctly (Mulligan, Hastedt, & McCarroll, 2012; Mulligan, McCarroll, Flanagan, & Potter, 2014; Mulligan, McCarroll, Flanagan, & Potter, 2015; Mulligan, McCarroll, Flanagan, & Potter, 2016; Tourangeau, et. al 2017). If answers on several easy items are wrong, the probability of a correct answer on a difficult item would be quite low. Omitted items are also less likely to cause distortion of scores, as long as enough items have been answered to establish a consistent pattern of right and wrong answers. The IRT-based overall scale score for math and science domains is an estimate of the number of items a child would have answered correctly in each data collection round if he or she had been administered all of the questions for that domain in all of the kindergarten, first grade, second grade, and third grade rounds. To calculate the IRT-based overall scale score for each domain, a child's IRT ability estimate (theta) is used to predict a probability for each assessment time that the child would have gotten that item correct. Then, the probabilities for all the items fielded as part of the domain in every round are summed to create the overall scale score. Because the computed scale scores are sums of probabilities, the scores are not integers. Finally, the IRT scoring makes possible longitudinal measurement of gain in achievement, even when assessments that are administered to a child are not identical at each point, for example, when a child was administered different levels of the second-stage form in the fall and spring data collections within one year or different sets of items across grades. To calculate the IRT-based overall scale score for each domain, a child's theta is use to predict a probability for each assessment item that the child would have gotten that item correct. Then, the probabilities for all the items fielded as part of the domain in every round are summed to create the overall scale score. Because the computed scale scores are sums of probabilities, the scores are not integers. Table 7 describes the region or locality of each school, which is not identified in the publicly available dataset as well as the school type (public).

Science Achievement. The science assessment domain included questions about physical science, life sciences, environmental sciences, and scientific inquiry. The science assessment included 19 routing items that all children received, followed by one of three second-stage forms (low, middle, or high difficulty). As with reading and mathematics, the second-stage form children received depended on their responses to the routing items (Tourangeau et al., 2017). The questions, response options, and any text the children could see on the easel pages (for example, graph labels) were read to the children to reduce the likelihood that their reading ability would affect their science assessment score. Kindergarten science knowledge and skills were measured using a 20-item assessment that was administered only in the spring data collection. All students were administered the entire assessment. A two-stage design was not needed for science because the length of the test was relatively short with respect to both time (approximately 10 minutes) and number of items.

Mathematics Achievement. Mathematics assessments used in all studies measured conceptual knowledge, procedural knowledge, and problem solving through items related to number properties, operations, geometry and spatial sense, data analysis, statistics, probability, patterns, algebra, and functions. For the mathematics assessment, most text that the children could see on the easel pages (for example, question text for word problems or graph labels) was read to the children to reduce the likelihood that their reading ability would impact their mathematics assessment performance. Spanish-speaking students who did not pass the language screener were administered the mathematics assessment that had been fully translated into Spanish. The mathematics assessment was not administered to students whose home language was one other than English or Spanish who did not achieve at least the minimum score on the screener. The possible range of scores was 0 to 96.

The cognitive assessments were individually administered by trained assessors using computer-assisted technology and small easel test books containing the assessment items. The reading and mathematics assessments were administered in both the fall and spring data collections using two-stage adaptive tests. For each assessment, the first-stage was a routing section that included items covering a broad range of difficulty. A child's performance on the routing section determined which one of the three second-stage tests (low, middle, or high difficulty) the child was administered. The second-stage tests varied by level of difficulty so that a child would be administered questions appropriate to his or her demonstrated level of ability for each of these cognitive domains. The purpose of this adaptive assessment design was to maximize accuracy of measurement while minimizing administration time. Table 8 provides the names of the variables pertaining to the IRT scale scores available in the data file, along with the variable descriptions, value ranges, weighted means, and standard deviations. Data includes fall 2010 math IRT scale score from kinder through second grade, spring 2011 math IRT scale scores from kinder through second grade, fall 2011 math IRT scale scores kinder through second, spring 2012 math IRT scale scores from kinder through second grade. Data also includes spring 2011 science IRT scale scores from kinder through second grade, fall 2011 science IRT scale scores kinder through second, spring 2012 science IRT scale scores from kinder through second grade, and fall 2012 science IRT scale scores.

Table 8
Cognitive Assessment Scale Scores from 2010-2013

Variable	Description	N	Range of possible values	Weighed mean	Standard Deviation
X1MSCALK2	F 2010 Math IRT Scale Score K-2	15,595	0.0-113.0	31.32	11.243
X2MSCALK2	S 2011 Math IRT Scale Score K-2	17,143	0.0-113.0	44.86	12.217
X3MSCALK2	F 2011 Math IRT Scale Score K-2	5,222	0.0-113.0	53.35	14.660
X4MSCALK2	S 2012 Math IRT Scale Score K-2	15,103	0.0-113.0	66.82	15.187
X2SSCALK2	S 2011 Science Scale Score K-2	16,936	0.0-64.0	28.07	7.526
X3SSCALK2	F 2011 Science Scale Score K-2	5,180	0.0-64.0	31.09	8.653
X4SSCALK2	S 2012 Science Scale Score K-2	15,072	0.0-64.0	36.29	9.198
X5SSCALK2	F 2012 Science Scale Score K-2	4,724	0.0-64.0	39.32	8.782

Reliability

Reliability statistics assess consistency of measurement, or the extent to which test items in a set are related to each other and to the score scale (Mulligan, McCarroll, Flanagan, & Potter, 2014; Tourangeau et al., 2017). Score reliability values range from 0 to 1. Table 8 presents the reliability statistics computed for the IRT-based scores for each subject area for the fall and spring of kindergarten, the fall and spring of first grade, and the fall of second grade. The reliability of the overall ability estimate, theta, is based on the variance of repeated estimates of theta for each individual child compared with total sample variance (Tourangeau et al., 2017). The reliabilities are relatively high, ranging from .75 to .95. According to Table 8, science is the domain with the most diverse content and the smallest number of items had lower reliability coefficients than mathematics.

Table 9
Reliability of IRT-Based Scores 2010-2013

Domain	Number of Items	Fall K/ F 2010	Spring K/ S 2011	Fall 1st F 2011	Spring 1st S 2012	Fall 2nd F 2012
Mathematics	113	.92	.94	.93	.93	.92
Science	64	NA	.75	.83	.83	.83

U.S. Department of Education, Early Childhood Longitudinal Study, National Center for Education Statistics, ECLS-K:2011, fall 2010, spring 2011, fall 2011, spring 2012, and fall 2012 (2017).

Executive Function and Achievement

Executive functions are interdependent processes that work together to regulate and orchestrate cognition, emotion and behavior and aide child learning in the classroom. Two measures of executive function were included in the kindergarten, first-grade, and second-grade assessment battery: the *Dimensional Change Card* (DCCS) (Zelazo, 2006) (Zelazo et al 2013), assessing children's cognitive flexibility or shifting (CS), and the Numbers Reversed subtest of the *Woodcock-Johnson III* (WJ III) *Tests of Cognitive*

Abilities (Woodcock, McGrew, and Mather 2001), assessing working memory (WM). The same versions of the DCCS and the Numbers Reversed tasks were administered in fall and spring of first grade. In second grade, the DCCS was changed to computerized administration to remain age appropriate, while the version of the Numbers Reversed task remained the same version used in the earlier data collection rounds (Tourangeau, et al., 2017). The variable names, descriptions, total values, value ranges, weighted means, and standard deviations for both the Dimensional Card Sort Game and the Numbers Reversed scores from the fall of kindergarten to the spring of second grade are shown in Table 10 below. The variables X1DCCSTOT refers to the card sort sort task combined score at time 1. The variables X2DCCSTOT refers to the card sort sort task combined score at time 2. The variables X2DCCSTOT refers to the card sort sort task combined score at time 2. The variables X3DCCSTOT refers to the card sort sort task combined score at time 3. The variables X4DCCSTOT refers to the card sort sort task combined score at time 4. The variable X1NRWABL refers to the Numbers Reversed task at time 1. The variable X2NRWABL refers to the Numbers Reversed task at time 2. The variable X3NRWABL refers to the Numbers Reversed task at time 3. The variable X4NRWABL refers to the Numbers Reversed task at time 4. The timing of the assessments includes the following: Fall 2010 card sort combined score, spring 2011 card sort combined score, fall 2011 card sort combined score, spring 2012 card sort combined score, fall 2010 numbers reversed score, spring 2011 numbers reversed score, fall 2011 numbers reversed score, spring 2012 numbers reversed score.

Table 10

Executive Function Variable Descriptors 2010-2012

Variable Name	Description	*n*	Range of possible values	Weighted Mean	Standard Deviation
X1DCCSTOT	F 2010 Card Sort Combined Score	15,604	0-18	14.18	3.343
X2DCCSTOT	S 2011 Card Sort Combined Score	17,149	0-18	15.14	2.815
X3DCCSTOT	F 2011 Card Sort Combined Score	5,222	0-18	15.89	2.293
X4DCCSTOT	S 2012 Card Sort Combined Score	15,109	0-18	16.05	2.347
X1NRWABL	F 2010 Numbers Reversed W-Ability Score	15,598	393-603	432.56	30.028
X2NRWABL	S 2011 Numbers Reversed W-Ability Score	17,147	393-603	449.49	30.412
X3NRWABL	F 2011 Numbers Reversed W-Ability Score	5,222	393-603	458.42	27.990
X4NRWABL	S 2012 Numbers Reversed W-Ability Score	15,107	393-603	459.56	25.395

Executive function (Dimensional Card Sort and Numbers Reversed)

The executive function component of the cognitive assessment obtained information on cognitive processes associated with learning: cognitive flexibility and working memory. To measure cognitive flexibility, children were administered the Dimensional Change Card Sort (DCCS) (Zelazo 2006). Different versions of the DCCS were used in different rounds of data collection because there was no single task that was age appropriate across all rounds of data collection when the study began. During the kindergarten and first-grade rounds, the hard-copy, tabletop version of the DCCS, as described in Zelazo (2006), was administered using physical cards that children were asked to sort into piles. Because the tabletop version of the DCCS would have been too easy for the majority of the study children during the second-grade rounds, in both the fall and the spring children were administered a new, age-appropriate, computerized version of the DCCS in which the "cards" are presented on a computer screen using keys to indicate where to place each card. The computerized task was developed as part of the National Institutes of Health (NIH) Toolbox for the Assessment of Neurological and Behavioral Function and is appropriate for ages 3-85 (Zelazo, Anderson, Richler, Waller-Allen, Beaumont, & Weintraub, 2013). The Toolbox DCCS has two versions that differ based on the age of the child: one version for children 7 years and younger and one for children 8 years and older. The ECLS-K:2011 used the version for children 8 years and older. Although the construct assessed in the tabletop and the computer versions is the same, the scoring and the way in which the construct is assessed differ across the two tasks.

Using scoring rules provided by the developers, four scores were developed from the DCCS data for the fall and spring kindergarten and the fall and spring first-grade rounds of data collection: the pre-switch score, the post-switch score, the Border Game score, and a total score. A final combined scale score reflects the total accuracy for the three tasks (i.e., the total number of cards sorted in the Color, Shape, and Border Games) which results in maximum score of 18 correct. The developer of the DCCS recommends using the overall accuracy score to assess performance (Tourangeau et al., 2017). The total scores for kindergarten and first grade (X1DCCSTOT, X2DCCSTOT, X3DCCSTOT, X4DCCSTOT) included in the kindergarten, first-grade, and second grade data files reflect children's performance across all 18 trials.

The construct assessed in the physical version and the computerized version of the DCCS is the same-cognitive flexibility explained within Chapter 2 of this proposed study.

Numbers Reversed

The Numbers Reversed task assess the child's working memory. It is a backward digit span task that requires the child to repeat an orally presented sequence of numbers in the reverse order in which the numbers are presented (Tourangeau et al., 2017). For example, if presented with the sequence "3…5," the child would be expected to say "5….3". Children are given 5 two-number sequences. If the child gets three consecutive two number sequences incorrect, then the Numbers Reversed task ends. If the child does not get three consecutive two number sequences incorrect, the child is then given 5 three number sequences. The sequence becomes increasingly longer, up to a maximum of eight numbers, until the child gets three consecutive number sequences incorrect or completes all number sequences.

Item-level data for the Numbers Reversed subtask for all and spring kindergarten, first grade, and second-grade are provided in the ECLS-K:2011 K-2 data file. Before analyzing the Numbers Reversed data, it is important that researchers understand the characteristics of these scores and how these characteristics may affect the analysis and interpretation of the Numbers Reversed data in the context of the ECLS-K:2011 (Tourangeau et. al, 2017). The *W* score may be best for a longitudinal analysis because it a measurement of growth and can be considered a growth scale (Tourangeau et al., 2017). The *W* score is a type of standardized score, is a type of transformation of the Rasch ability scale, and provides a common scale of equal intervals that represents both a child's ability and the task difficulty. Typically, it has a mean of 500 and standard deviation of 100. Most children younger than 10 years old would obtain W scores lower than the mean of 500, and most older children would be expected to have scores above the mean of 500. As a child develops with age, it would be expected that the child's *W* score would increase to reflect growth. Researchers and readers should keep in mind that most ECLS-K:2011 sample children were 5 or 6 years old during the kindergarten data collections, 6 or 7 years old during the first-grade data collections, and 7 or 8 years old during the second-grade data collections while the W scores compare their performance to that of 10-year-olds. As a result, *W* scores from the ECLS-K:2011 sample appear to show that the ECLS-K:2011 children demonstrated below average performance on this task. As the children grow older and closer to the age of 10, the discrepancy declines.

Procedure for Testing Hypotheses and Answering Research Questions

Structural equation modeling provides powerful framework for investigating complex factors affecting scores on educational achievement tests over time. Using the Early Childhood Longitudinal Study Cohort 2011 (ECLS-K:2011), we leverage SEM's flexible framework to address our research goals given the complex data structure (e.g., individually varying time points of measurement, missing data elements, hierarchical structures, and non-constant variability across achievement scores).

Structural Equation Modeling (SEM)

Structural equation modeling provides a powerful framework for testing *apriori* hypotheses about a variety of causal models. Structural equation modeling (SEM) provides a statistical framework that allows for a set of relationships between one or more independent variables and one or more dependent variables. In fact, CFA is a type of SEM that deals specifically with measurement models, that is the relationships between observed measures or indicators (test items, test scores or behavioral ratings) and latent variables or factors (Brown, 2006). The independent and dependent variables may be latent or observed variables and the level of measurement may be discrete or continuous. SEM is also known as causal modeling, covariance structure modeling, or simultaneous equation modeling. Considering that variables in achievement cannot be measured directly, they can be accounted for through the measurement of certain observable variables that define or are though to define them. Since the use of latent variables enables errors in such variables to be identified, estimated values of variables in SEM studies are much more reliable (Simsek, 2007). SEM is an approach used to test the models characterized by causal and correlational relationships between observable and latent variables, and it allows one to study the set of relationships between observable and latent variables, and it allows one to study the set of relationships between one or more independent variables and one or more dependent variables (Anagun, 2011).

We employ a parallel process latent growth model with covariates which allows us to estimate latent means, variances and covariances necessary to address our research issues provided in section I. To evaluate if the PPLGM is adequately powered, we applied recommended procedure of MacCallum, et. al (1996). We will conduct a Monte Carlo simulation in M*plus* based on the PPLGM model to ensure that adequate power would exist for all estimated model parameters. In both instances, adequate power for the PPLGM will be verified based on the sample size available to us in the ECLS-K:2011.

Prior to all analyses and across all time points, data screening was conducted and included identifying and/or evaluating missing data elements, univariate and multivariate outliers, normality, heteroscedastic variance and linearity. Score reliability and measurement invariance for the outcome variables were evaluated (Grimm, et. al, 2017). Missing data values were replaced using full-information maximum likelihood (FIML) estimation within the M*plus* program. Trajectories for both mathematics and science scores were observed to be linear providing support for a linear approach for our growth models. Finally, the ECLS-K:2011 full sample weights will be included in all analyses to ensure accurate standard errors of parameter estimates.

Summary

Scores on tests of educational achievement may be influenced by two or more co-occurring variables (a.k.a. covariates). In this study, co-occurring variables related to educational achievement in mathematics and science include Spanish spoken at home (i.e. English as a second language rather than primary) and assessments of executive function. The existence of co-occurring factors and their relationship to how students' mathematics and science achievement scores change over time presents a substantial challenge to providing effective student-centered interventions.

We introduce an analytic model with the goal of increasing understanding about the complex relations between student-level covariates and achievement scores using parallel process latent growth model (PPLGM) with covariates (Figure 1.). Specifically, we examine how growth trajectories of children's mathematics and science scores change over time in while including student-level covariates. The present study was based on structural equation modeling and employed a population-based and multi-year longitudinal sample of U.S. schoolchildren in public schools, the ECLS-K:2011 (NCES, 2011) and focused on kindergarten, first grade, and second grade.

Prior to conducting secondary all analyses, data were screened for missing data and univariate and multivariate outliers using Mahalanobis distance and boxplots. Data were also screened to ensure the assumptions of normality were tenable using histograms and other tests of normality such as normal Q-Q plots, linearity using scatterplots, and looking at residual plots for homoscedasticity. Next, structural equation modeling (SEM) was employed using *M*plus version 7.4 to provide a statistical framework that allows a set of relationships to exist between one or more independent variables and one or more dependent variables.

Intervention strategies that improve domain general executive function skills in children whose economic background puts them at risk for school failure could have lifelong benefits for academic success.

IV. FINDINGS

Data Screening

Prior to conducting secondary data analyses, data were screened beforehand for missing data and univariate and multivariate outliers using Mahalanobis distance and boxplots. Data were also screened to ensure the assumption of normality was tenable using histograms and other tests of normality such as normal Q-Q plots, linearity using scatterplots, and looking at residual plots for homoscedasticity are met. Following linear regression modeling, structural equation modeling (SEM) was used with the *Mp*lus program version 7 to provide a statistical framework that allows for a set of relationships between one or more independent variables and one or more dependent variables. The strength of SEM is that is allows both confirmatory factor analysis for measurement models and path analysis for latent variable models to be processed simultaneously.

The data were also screened for univariate normality. Suggested cutoffs of skewness < |2.0| and kurtosis > |7.0| as indicated by West, Finch, and Curran (1995). "Based on the results of data screening, the data distributions for the manifest variables in this study exhibited univariate and multivariate kurtosis. To accommodate for this artifact, the maximum likelihood robust (MLR) estimator was used to derive parameter estimates in all analyses." (Muthén & Muthén 1998-2014).

Missing Values

The pattern of missing data is more important than the amount missing. Missing values scattered throughout the data are preferable to non-randomly missing values because they affect the generalizability of the results. Data can be missing at random

(MAR) or missing completely at random (MCAR). If only 5% or less are missing in a random pattern from a large data set, then a procedure for elimination of data will not affect the result. The procedure for handling missing data in this proposed study is

Outliers

Univariate outliers can be visually detected by inspecting the data, or by the use of boxplots where extreme values are located far away from the box. A check for univariate statistics and Mahalanobis distance were obtained from the *M*plus 7 PLOT command (Muthén & Muthén 1998-2014) and did not display any extreme outliers. For multivariate outliers, Mahalanobis distance is evaluated as a chi square ($\chi 2$) statistic with degrees of freedom equal to the number of variables in the analysis (Mertler and Vannatta, 2013). The accepted criterion for outliers is a value for Mahalanobis distance that is significant beyond $p<.001$. Another way to assess outliers is to use SPSS DESCRIPTIVES and deleting the cases that have extremely high or extremely low scores prior to factor analysis.

Measurement Errors

The measurement error terms are independent of each other and of the factors. All associations between the factors are unanalyzed. *E*'s are the measurement errors (represented by the small circles in Figure 1), which represent the unique variance, a factor analytic term for indicator variance unexplained by the factors. Measurement errors are proxy variables for all sources of residual variance that are not explicitly represented in the model. Two types of unique variance are presented by *E* terms: random error (unreliability) and all sources of systematic (nonrandom) variance not due to the factors. Note that each subtest in this model is specified to measure (load on) a single latent variable. In contrast, standard EFA does not permit the researcher to specify that a subtest loads on only one factor. Factor loadings in CFA are more generally interpreted as regression coefficients in unstandardized and standardized form. The line that points to an indicator from its measurement error term (Figure 1) reflects all other sources of variance not explained by the indicator's underlying factor. Measurement error terms reflect two kinds of unique variance: 1) random error of the type estimated by reliability coefficients, and 2) systematic variance due to things that the indicator measures besides its underlying factor.

Maximum Likelihood Estimation

Maximum likelihood (ML) estimation is the most frequently used method for structural equation modeling. It is the default in most SEM computer programs, and most structural equation models described in the literature are analyzed with this method (Kline, 2011). The term maximum likelihood describes the statistical principle that underlies the derivation of parameter estimates; the estimates are the ones that maximize the likelihood (the continuous generalization) that the data (the observed covariances) were drawn from this population. Most forms of ML estimation in SEM are simultaneous, which means that the estimates of model parameters are calculated all at once.

Use of the *M*plus Program

*M*plus is a statistical modeling program that provides researchers with a flexible tool to analyze their data by offering choice of models, estimators, and algorithms in program that has an easy to use interface and graphical displays of data and analysis results (Muthén & Muthén, 2015). Mplus allows the analysis of both cross-sectional and longitudinal data, single-level, and multilevel data, data that come from different populations with either observed or unobserved heterogeneity, and data that contain missing values. Analyses can be carried out for observed variables that are continuous, censored, binary, ordered categorical (ordinal), unordered categorical (nominal), counts, or combinations of these variable types. The *M*plus modeling framework draws on the unifying theme of latent variable modeling where measurement of unobserved constructs is realized via unified measurement and structural components.

Population and Sample Participants

Descriptive statistics were conducted using SPSS version 24.0 using the publicly available ECLS-K:2011 dataset downloaded via the ECLS-K:2011 electronic codebook. The original ECLS-K:2011 data consists of 57,000 variables and a sample size of 18,174 participants. Demographics and descriptive data for the full sample ECLS-K:2011 tables include gender, child birth year, ethnicity, school location, school type, and income category. Of the original sample of 18,174 children in the original dataset (see Table 11 and Table 12), an unweighted sample of 2,272 (see Table 13) was filtered and limited to those students attending public schools in city and suburban areas. The unweighted sample consisted pf 41.8% Hispanic, 30.8% White, 9.8% African American, and 9.2%

Asian students ages 4 and 5 years throughout the four time points. For the time-invariant, binary Spanish in the home variable, those answering "Yes" to the question constituted 21.5%, 22.4%, 22,7%, and 23.0% respectively, throughout the four time points in the unweighted sample. Demographic and descriptive data for the unweighted sample included ethnicity, school location, school type, and Spanish of the home (Yes, No). Table 12 describes the demographic sample by school location (city, suburb, town, rural) by the total number of participants in each sample, the income category, and the school type (public and private). Most of the students in the entire sample ECLS-K: 2011 are found in $75,000 to $200,000 income category.

Table 11

Demographics of the Full ECLS-K:2011 Gender and Ethnicity

Gender	*n*	Percent	Child DOB Year	Ethnicity	*n*	Percent
Female	8849	48.7	2003/2004	American Indian/Alaska Native	169	0.9
Male	9283	51.1	2005/2006	Asian, non-Hispanic	1546	8.5
Not ascertained	42	0.2	Not ascertained	Black/African American, non-Hispanic	2397	13.2
				Hispanic, No Race Specified	641	3.5
				Hispanic, Race Specified	3944	21.7
				Native Hawaiian/Pacific Islander, non-Hispanic	116	0.6
				Not ascertained	50	0.3
				Two or more races, non-Hispanic	822	4.5
				White, not Hispanic	8489	46.7
Total	18174	100.0			18174	99.9

Table 12

Demographics of the Full ECLS-K:2011 Income Category

School Location	*n*	Income Category	*n*	School Type	*n*
Time 1		$5,000 or less	408	Time 1	
City	5382	$5,000 to $10,000	562	Public	14329
Suburb	5757	$10,000 to $15,000	845	Private	2135
Town	1317	$15,000 to $20,000	916	Time 2	
Rural	3751	$20,000 to $25,000	1047	Public	15602
Time 2		$25,000 to $30,000	636	Private	2189
City	5963	$30,000 to $35,000	668	Time 3	
Suburb	6340	$35,000 to $40,000	636	Public	4900
Town	1337	$40,000 to $45,000	455	Private	434
Rural	3885	$45,000 to $50,000	492	Time 4	
Time 3		$50,000 to $55,000	430	Public	4900
City	2292	$55,000 to $60,000	417	Private	434
Suburb	1991	$60,000 to $65,000	445		
Town	212	$65,000 to $70,000	391		
Rural	760	$70,000 to $75,000	511		
Time 4		$75,000 to $100,000	1764		
City	5066	$100,000 to $200,000	2257		
Suburb	5444	$200,000 or more	588		
Town	1200	not ascertained	81		
Rural	3264	Missing	4563		
Total			18112		

Table 13

Demographic Composition of the Unweighted Sample

Ethnicity	*n*	Percent	School Location	*n*	School Type	*n*	Spanish Home	*n*	Percent
American Indian/Alaska Native	24	0.9	Time 1		Time 1		Time 1		
Asian	255	9.2	Suburb	1373	Public	2768	Yes	595	21.5
Black/African American	271	9.8	City	1399	Time 2		No	2025	73.1
Hispanic, Race Specified	1160	41.8	Time 2		Public	2772	Missing	152	5.5
Hispanic, No Race Specified	109	3.9	Suburb	1374	Time 3		Time 2		
Native Hawaiian.Pacific Islander	14	0.5	City	1398	Public	2772	Yes	622	22.4
Not Ascertained	3	0.1	Time 3		Time 4		No	2096	75.6
Two or More Races	97	3.5	Suburb	1332	Public	2772	Missing	54	1.9
White	839	30.3	City	1440			Time 3		
			Time 4				Yes	629	22.7
			Suburb	1338			No	2107	76
			City	1434			Missing	36	1.3
							Time 4		
							Yes	637	23
							No	2110	76.1
							Missing	25	0.9
Total	2772	100.0		2772				2747	100

After the full sample weights were applied, the new sample size became 629,706 students and now generalized to a much larger population to compensate for the unequal probabilities of selection and non-response at each sampling stage. Replicate weights were added which were designed for the analysis of child direct assessments. After the full sample weights were applied, the new sample size became 629,706 students and now generalized to a much larger population (see Table 14). The weighted sample consisted of 42.9% Hispanic, 32.9% White, 9.9% African American, and 5.5% Asian students located in public schools in suburbs and cities throughout the four time points. For the time-invariant, binary Spanish in the home variable, those answering “Yes” to the question constituted 23.7%, 23.5%, 23.7%, and 23.8% respectively, throughout the four time points in the weighted, more representative sample.

Table 14

Demographic composition of the Weighted Samples

Ethnicity	*n*	Percent	School Location	*n*	School Type	*n*	Spanish Home	*n*	Percent
American Indian/Alaska Native	5663	0.8	Time 1		Time 1		Time 1		
Asian	37463	5.4	Suburb	334640	Public	691569	Yes	164491	23.7
Black/African American	68755	9.9	City	358066	Time 2		No	524222	75.7
Hispanic, No Race Specified	29920	4.3	Time 2		Public	692706	Time 2		
Hispanic, Race Specified	296930	42.9	Suburb	334508	Time 3		Yes	162875	23.5
Native Hawaiian/Pacific Islander	2202	0.3	City	358198	Public	692706	No	522055	75.4
Not ascertained	854	0.1	Time 3		Time 4		Time 3		
Two or more races	23133	3.3	Suburb	326194	Public	692706	Yes	163234	23.6
White	227787	32.9	City	366512			No	522687	75.5
			Time 4				Time 4		
			Suburb	327673			Yes	164533	23.8
			City	365033			No	522963	75.5
Total	692707	99.9		692706		692706		687496	99.3

Estimating Longitudinal Growth Curve Model

First, the estimation of the longitudinal covariance patterns of the observed variables, math and science achievement scores over four time points was performed to ensure a longitudinal growth curve model was feasible. At stage 1, the goal was to examine individual student change over time (see Figure 13 and Figure 14). Each student has a regression line (growth curve) that plots the variable of interest over time. An unconditional (with no covariates) parallel process longitudinal growth curve model was estimated to analyze change over time in the two outcomes of interest: mathematics and science achievement (see Figure 13 and Figure 14 below). Linear and non-linear models were tested to find the best fit for the observed data patterns. Modification indices were also examined to improve model fit. Unconditional LGCMs include two latent factors represented by ovals (Figures 13-15 below), the intercept and the sloped and repeated measures of the observed outcomes of interest over time (represented by rectangles in Figures 13-15 below). The latent growth factors were assumed to covary. To test whether the growth parameters of one curve was associated with the growth parameters of the other, the best-fitting models were then combined into one unconditional parallel process model (see Figure 14).

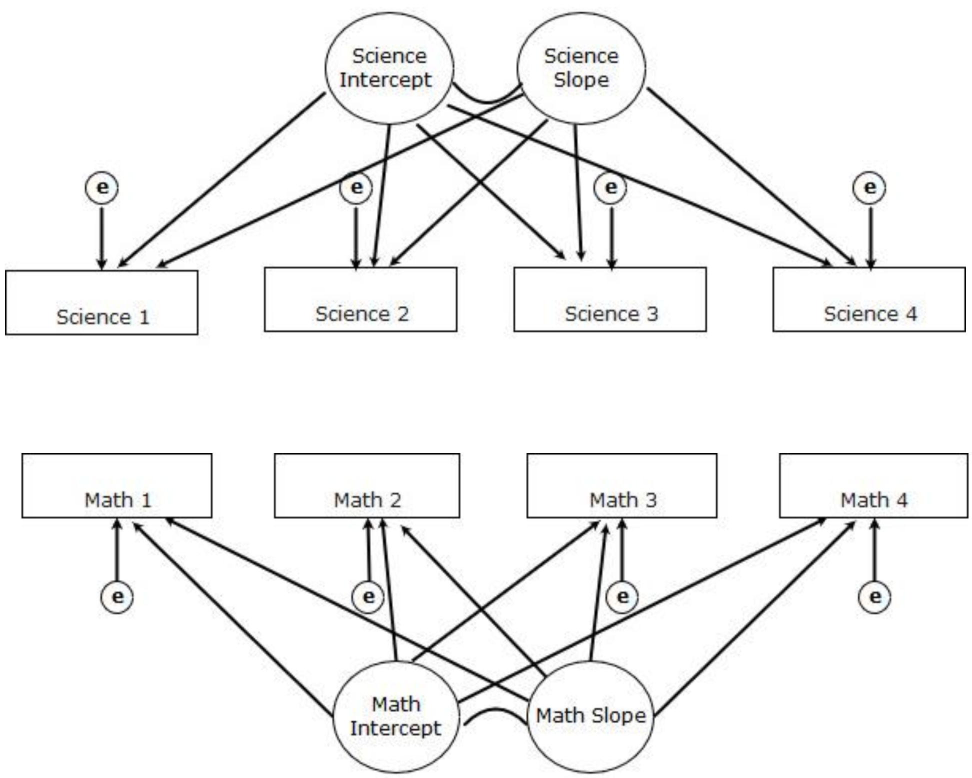

Figure 13. *Individual Unconditional Univariate Math and Science Model*. Latent constructs are shown in ellipses, and observed variables are shown in rectangles. Intercepts and slopes covary.

At the second stage the individual level regression equations can then be summarized to obtain an average intercept (or initial level) and average slope, or average rate of change for all students (see Figure 14). The intercept and slope each have their own variance, and as such, this is how much "variability" there is for student scores at each time point (variance intercept) and over time (slope variance). The first model analyzed is the unconditional multivariate model where math and science constitute the observed variables and no covariates are present (see Figure 14 below). Unconditional LGCM include two latent (represented by ovals in Figures 14-15).

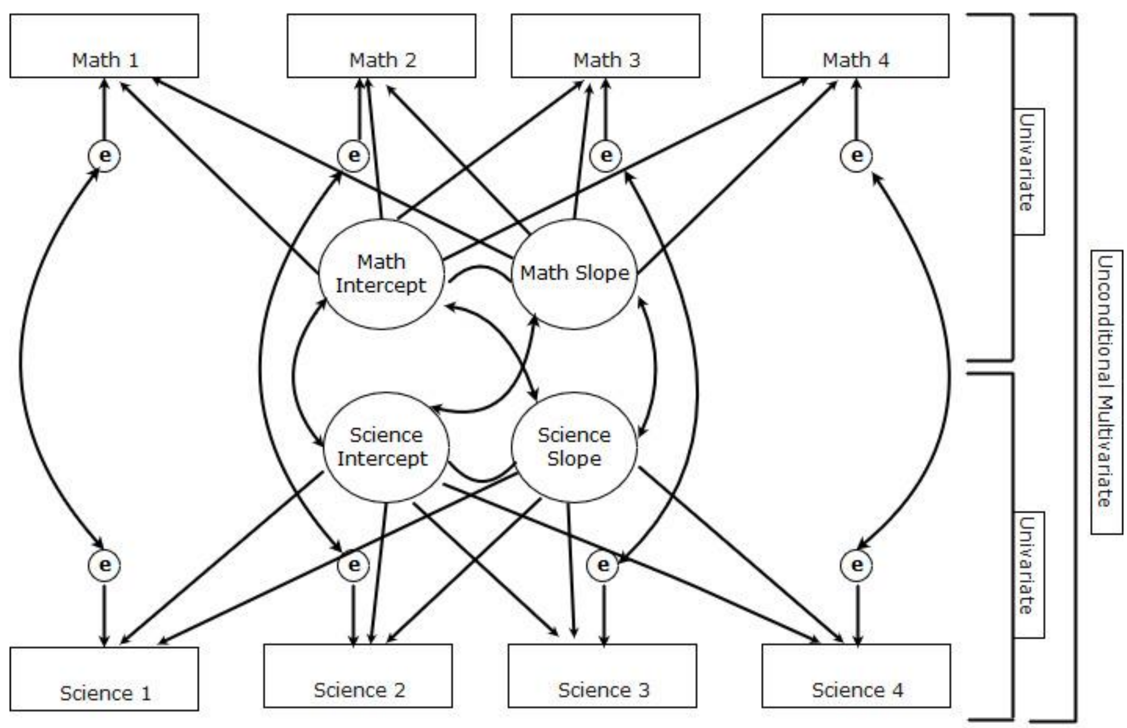

Figure 14. *Unconditional Multivariate Model*. Latent constructs are shown in ellipses, and observed variables are shown in rectangles.

The last step in estimating the models involves the multivariate conditional model (by adding covariates) that was estimated to examine the associations between growth parameters (intercepts and slopes) and antecedents or distal outcomes in the full parallel process latent growth curve model. The intercept and slope growth factors were regressed on three time-varying and one time-invariant predictor(s): two measures of executive function, Numbers Reversed task and Card Sort task, as well as Spanish spoken in the home.

All steps involved in the model analyses are summarized in detail in Figure 16. The covariates added were two measures of executive function, Numbers Reversed and Card Sort, as well as Spanish spoken in the home. The full information maximum-likelihood estimator was used to account for missing data, signifying that all observations from the dataset were used, including participants with missing data at one or more waves of data collection. To account for non-normality and skewness of variables, all models used maximum likelihood estimation with robust standard errors (Muthén & Muthén 1998-2014).

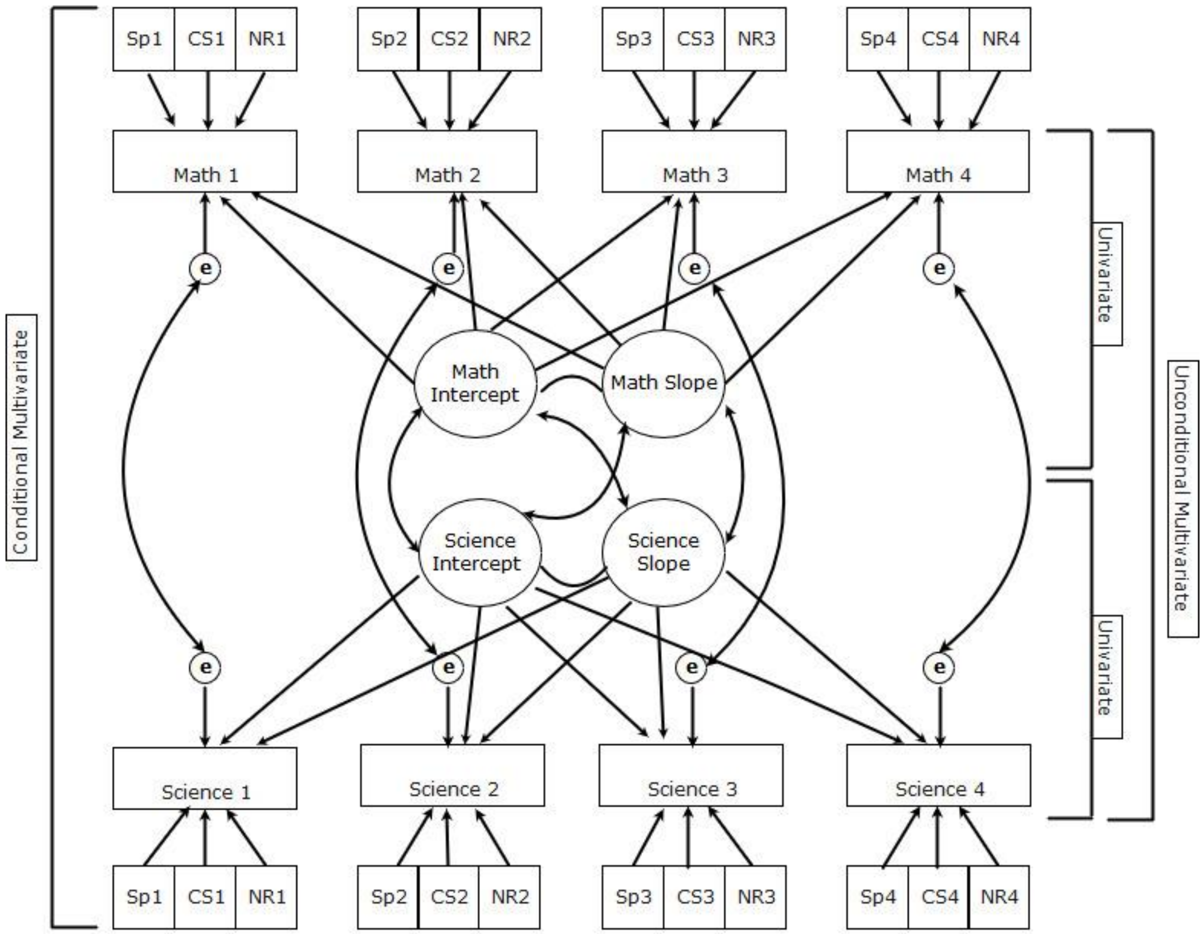

Figure 15. *Conditional Multivariate Model*. Latent constructs are shown in ellipses, and observed variables are shown in rectangles.

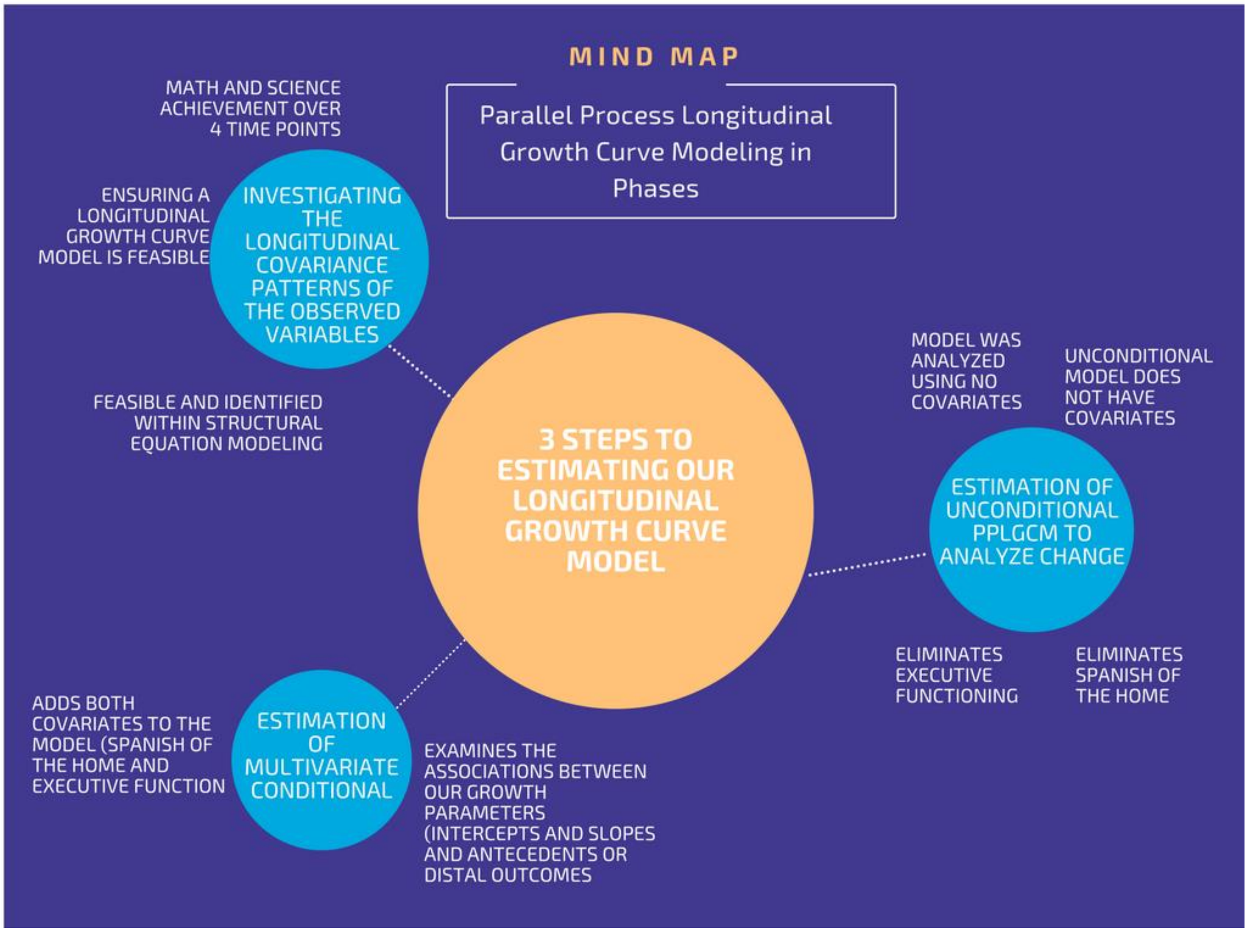

Figure 16. *Mind Map Outlining the Growth Curve Model in Three Steps*

Model Fit Indices

Goodness-of-fit values for the latent variables used in the model were assessed according to present standards in the literature. As measures of model fit, the $\chi 2$ goodness-of-fit test, the root mean square error of approximation (RMSEA), (TLI), and the comparative fit index (CFI) were used. Hu and Bentler (1999) suggested that for continuous data, RMSEA < .06, TLI > .95, CFI > .95. The RMSEA is an established tool for evaluating model fit as it takes both the number of observations and the number of free parameters into account (Brown, 2006). Browne and Cudeck (1993) suggest that values below .05 indicate a close fit and values between .05 and .08 indicate an acceptable model fit, while Hu and Bentler (1998) suggest that values below .06 indicate a good model fit. The Root Mean Square Error of Approximation (RMSEA) is scaled as a badness-of-fit index where a value of zero indicates the best fit. The value of RMSEA decreases as there are more degrees of freedom (greater parsimony) or a larger sample size, which is the observed in table 15, where the multivariate conditional model has 112 degrees of freedom (see Table 15). For purposes of the present study (see Table 15) the multivariate conditional model (i.e. the PPLGCM) displayed an RMSEA value of 0.08 indicating an acceptable level of fit (i.e. a value closer to zero indicates a better fitting model). All the other models in Table 15 display an RMSEA value of 0.404, 0.114 and 0.22 exceeding the recommended values between 0.05 and 0.08. The Comparative Fit Index (CFI) is an incremental fit index that measures the relative improvement in the fit of the researcher's model over that of a baseline model, typically the independence (completely uncorrelated) model. CFI compares the χ^2 value of the model to the χ^2 of the null model. Values can range from 0 to 1; however, values larger than or equal to .95 are indicative of a good model fit given the observed data (Brown, 2006). Such is the case in the present study, where the conditional multivariate model (with covariates) has a value of 0.905 and the unconditional multivariate model (no covariates) has a value of 0.918, both indicating adequate fit. The SRMR (standardized root mean square residual) allows one to evaluate the quality of model-data fit by examining the residuals (i.e. residual being the discrepancy between the proposed model and the covariance matrix actually used based on the sample data). For the SRMR, values ranging from 0 to 1 with well-fitting models less than 0.05 (Byrne, 2010). For the present study, the unconditional multivariate model has an SRMR value of 0.05 indicating a good fit for the model without covariates. A series of LGCM's were fit. The univariate LGCM's in mathematics and science were modeled first to assess model fit. The model specifications were adjusted as needed to improve model fit.

Table 15
Goodness-of-Fit Indicators of Six Models

Model	χ2	df	*p* value	χ2/df	RMSEA	90% CI	CFI	SRMR
Model 1 (UMNC)	37.09	2	0.000	18.54	0.40	0.38, 0.43	0.89	0.10
Model 2 (UMC)	2299.34	62	0.000	37.09	0.11	0.11, 0.12	0.81	0.26
	$\Delta \chi^2$ = 1392.46	Δdf = 60						
Model 3 (USNC)	906.88	2	0.000	453.44	0.40	0.38, 0.43	0.89	0.10
Model 4 (USC)	2299.34	62	0.000	37.09	0.11	0.11, 0.19	0.81	0.26
	$\Delta \chi^2$ = 1392.46	Δdf = 60						
Model 5 (MC)	2247.5	112	0.000	20.067	0.08	0.08, 0.09	0.90	0.21
Model 6 (MNC)	1551.52	12	0.000	129.29	0.22	0.20, 0.22	0.92	0.05
	$\Delta \chi^2$ = 695.98	Δdf = 100						

Note. UMNC=univariate math unconditional; UMC=univariate math conditional; USNC=univariate science unconditional

Chi-Square and Model Fit Indices for Present Study

First, it is helpful to assess the overall fit of the model by looking at the χ^2 and fit indices (see Table 15). Kline (2011) suggests that the chi-square test is overly sensitive to sample size when testing whether the same factor structure holds across different groups, that is, whether the measurement model is invariant over samples. When data are normally distributed and models are nested (one model is a subset of another), the χ^2 value for the larger model is subtracted from the χ^2 value for the smaller nested model and the difference, also considered a χ^2, is evaluated with degrees of freedom equal to the difference between the degrees of freedom in the two models (Tabachnick & Fidell, 2007). The factor loadings for the intercept latent variable were fixed to one.

Such is the case in the present study, where the conditional multivariate model (with covariates, see Table 15) has a value of 0.905 and the unconditional multivariate model (no covariates) has a value of 0.918. Model fit indices indicate adequate fit of the unconditional linear model ($\chi2$=1551.52, *df*=12, $p< 0.001$; CFI=0.918, RMSEA=0.215; 90% CI [0.206, 0.22], SRMR=0.05). For the present study, the unconditional multivariate model has an SRMR value of 0.05 indicating a good fit for the model without covariates. Model fit indices indicate good fit of the conditional linear model ($\chi2$=2247.5, *df*=112, $p< 0.001$; CFI=0.91, RMSEA=0.08; 90% CI [0.08, 0.09], SRMR=0.21). For purposes of the present study (see Table 15) the multivariate conditional model, also known as the PPLGCM has an RMSEA value of 0.08 indicates an acceptable fit because a value closer to zero indicates the best fit. The following indices were used to examine model fit: The CFI, TLI, RMSEA, and SRMR. Adequate fit was indicated by CFI and TLI > 0.90, RMSEA < 0.08, and SRMR < .10. Good fit was indicated by RMSEA < .06, TLI > .95, CFI > .95 and SRMR < .08 (Hu & Bentler, 1999).

Means, Standard Deviations, Skewness, and Kurtosis

The table below includes the descriptive summary of all the variables in the unweighted sample with mean, standard deviation, skewness and kurtosis.

Table 16

Descriptive Summary of Unweighted Variables

				Skewness		Kurtosis	
Model	N	Mean	Std. Deviation	Statistic	Std. Error	Statistic	Std. Error
Math 1 IRTSS	2620	29.9800	12.3389	0.292	0.048	1.289	0.096
Math 2 IRTSS	2718	43.2599	12.4111	-0.113	0.047	0.962	0.094
Math 3 IRTSS	2736	52.0554	14.7784	0.321	0.047	0.293	0.094
Math 4 IRTSS	2747	64.6158	15.5092	-0.265	0.047	0.699	0.093
Science 1 IRTSS	2718	25.0939	9.2945	-0.924	0.047	2.567	0.094
Science 2 IRTSS	2736	29.2274	9.4840	-0.109	0.047	1.269	0.094
Science 3 IRTSS	2747	34.1489	10.0803	-0.265	0.047	0.613	0.094
Science 4 IRTSS	2742	37.2140	9.8141	-0.587	0.047	0.330	0.093
nrscore 1	2620	423.68	57.965	-5.245	0.048	37.581	0.096
nrscore 2	2718	443.21	42.237	-4.943	0.047	50.849	0.094
nrscore 3	2736	455.42	33.930	-3.996	0.047	49.887	0.094
nrscore 4	2747	465.84	34.449	-6.016	0.047	77.178	0.093
cardsort 1	2620	13.60	4.329	-2.454	0.048	8.129	0.096
cardsort 2	2718	14.98	3.127	-3.016	0.047	14.880	0.094
cardsort 3	2736	15.57	2.596	-2.918	0.047	15.589	0.094
cardsort 4	2747	16.01	2.522	-3.574	0.047	23.641	0.093
sspeak 1	2620	1.77	0.419	-1.304	0.048	-0.301	0.096
sspeak 2	2718	1.77	0.420	-1.292	0.047	-0.332	0.094
sspeak 3	2736	1.77	0.421	-1.285	0.047	-0.350	0.094
sspeak 4	2747	1.77	0.422	-1.271	0.047	-0.384	0.093

The next table includes the descriptive summary of all the variables in the weighted sample with mean, standard deviation, skewness and kurtosis. As noted in Table 16, the mean for Math 1 IRT score, Math 2 IRT Score, Math 3 IRT score, Math 4 IRT score steadily increases. As the means in math scores increase, the means of Science 1 IRT score, Science 2 IRT score, Science 3 IRT score, and Science 4 IRT scores also increase. The same is true for the two time-varying covariates, numbers reversed (measuring cognitive flexibility) and cardsort game (measuring working memory). The only measure remaining constant is the time-invariant Spanish of the home throughout the four time points. The next table includes the descriptive summary of all the variables in the weighted sample with mean, standard deviation, skewness and kurtosis. After the addition of the full sample weights, the new sample size became 629,706 students and now generalized to a much larger population. The weighted sample shown in Table 17 depicts the mean for Math 1 IRT score, Math 2 IRT Score, Math 3 IRT score, Math 4 IRT score steadily increases. As the means in math scores increase, the means of Science 1 IRT score, Science 2 IRT score, Science 3 IRT score, and Science 4 IRT scores also increase. The same is true for the two time-varying covariates, numbers reversed (measuring cognitive flexibility) and cardsort game (measuring working memory). The only measure remaining constant is the time-invariant Spanish of the home throughout the four time points.

Table 17
Descriptive Summary of Weighted Variables

Model	N	Mean	Std. Deviation	Skewness Statistic	Skewness Standard Error	Kurtosis Statistic	Kurtosis Std. Error
Math 1 IRTSS	692706	29.94	11.31	0.719	0.003	0.812	0.01
Math 2 IRTSS	692706	43.09	11.98	0.129	0.003	0.062	0.01
Math 3 IRTSS	692706	51.86	14.5	0.424	0.003	0.008	0.01
Math 4 IRTSS	692706	64.46	15.02	-0.072	0.003	-0.157	0.01
Science 1 IRTSS	692706	25.78	7.9	0.083	0.003	-0.521	0.01
Science 2 IRTSS	692706	29.42	8.95	0.355	0.003	-0.25	0.01
Science 3 IRTSS	692706	34.30	9.68	0.039	0.003	-0.664	0.01
Science 4 IRTSS	692706	37.27	9.65	-0.481	0.003	-0.214	0.01
nrscore 1	692706	428.89	29.85	0.561	0.003	-1.051	0.01
nrscore 2	692706	445.00	30.89	-0.101	0.003	-1.118	0.01
nrscore 3	692706	456.20	28.83	-0.465	0.003	-0.208	0.01
nrscore 4	692706	466.96	26.08	-0.772	0.003	0.493	0.01
cardsort 1	692706	13.86	3.47	-1.393	0.003	0.882	0.01
cardsort 2	692706	15.07	2.75	-1.868	0.003	3.771	0.01
cardsort 3	692706	15.61	2.42	-2.056	0.003	5.843	0.01
cardsort 4	692706	16.07	2.22	-2.093	0.003	5.819	0.01
sspeak 1	692706	1.76	0.43	-1.217	0.003	-0.505	0.01
sspeak 2	692706	1.76	0.43	-1.227	0.003	-0.471	0.01
sspeak 3	692706	1.76	0.43	-1.224	0.003	-0.483	0.01
sspeak 4	692706	1.76	0.43	-1.209	0.003	-0.481	0.01

Table 18 shows the correlation coefficients of the achievement over time in mathematics and science at four different time points. Mathematics and science IRT scores were moderately correlated with each other (r's> 0.5) in all four time points. Consistent with the literature, the measure of executive function-the numbers reverse task measuring cognitive flexibility is highly correlated to mathematics achievement (r's> 0.5) in all four time points. Also consistent with the literature, the measure of the measure of executive function-the card sort task measuring working memory was moderately correlated to mathematics achievement (*r*’s> 0.3) in all four time points. As for as science achievement, and consistent with the literature, the measure of executive function-the numbers reverse task measuring cognitive flexibility is moderate to highly correlated to science achievement (*r*’s> 0.4) in all four time points. Consistent with the science literature, the measure of executive function-the card sort task measuring working memory is highly correlated to science achievement (*r*’s> 0.3) in all four time points. As for Spanish of the home and mathematics achievement, Spanish of the home has a weak correlation to mathematics achievement (*r*’s> 0.2) in all four time points. For science achievement, Spanish of the home throughout the four time points has a moderate correlation (*r*’s> 0.4) with science achievement.

Table 18
Correlations Among Weighted Variables

Variable	1	2	3	4	5	6	7	8
Achievement over time								
Math 1 IRTSS	1.00							
Math 2 IRTSS	0.805	1.00						
Math 3 IRTSS	0.774	0.844	1.00					
Math 4 IRTSS	0.758	0.822	0.846	1.00				
Science 1 IRTSS	0.581	0.602	0.576	0.602	1.00			
Science 2 IRTSS	0.618	0.640	0.639	0.644	0.816	1.00		
Science 3 IRTSS	0.609	0.641	0.629	0.680	0.794	0.854	1.00	
Science 4 IRTSS	0.601	0.647	0.639	0.681	0.768	0.818	0.86	1.00

Table 18
Continued

Predictors								
nrscore 1	0.634	0.592	0.581	0.587	0.509	0.520	0.520	0.502
nrscore 2	0.552	0.619	0.577	0.588	0.490	0.489	0.499	0.507
nrscore 3	0.510	0.559	0.586	0.567	0.411	0.460	0.453	0.479
nrscore 4	0.476	0.544	0.527	0.568	0.387	0.406	0.436	0.452
cardsort 1	0.357	0.357	0.341	0.358	0.339	0.354	0.350	0.360
cardsort 2	0.316	0.367	0.350	0.375	0.322	0.338	0.335	0.355
cardsort 3	0.305	0.349	0.356	0.351	0.297	0.329	0.332	0.346
cardsort 4	0.278	0.325	0.338	0.366	0.295	0.315	0.342	0.348
sspeak 1	0.278	0.267	0.242	0.272	0.459	0.407	0.406	0.374
sspeak 2	0.271	0.266	0.239	0.269	0.448	0.401	0.401	0.369
sspeak 3	0.276	0.264	0.244	0.271	0.451	0.399	0.403	0.369
sspeak 4	0.276	0.266	0.241	0.271	0.450	0.401	0.401	0.368

Interpretation of R^2 Results

Inspection of Table 19 reveals how much the three covariates contributed to the multivariate mathematics and science models. For Math 1 at Time 1, Spanish spoken in the home (SP1) accounted for 0.3% of the variance (R^2=.003, p< 0.001) in mathematics achievement. For Math 2 at Time 2, Spanish spoken in the home (SP2) accounted for 0.3% of the variance (R^2=.003, p< 0.001) in mathematics achievement. For Math 3 at Time 3, Spanish spoken in the home (SP3) accounted for 0.2% of the variance (R^2=.002, p< 0.001) in mathematics achievement. For Math 3 at Time 3, Spanish spoken in the home (SP3) accounted for 0.1% of the variance (R^2=.001, p< 0.001) in mathematics achievement. For Math 4 at Time 4, Spanish spoken in the home (SP4) accounted for 0.3% of the variance (R^2=.003, p< 0.001) in mathematics achievement. As for Science 1 at Time 1, Spanish spoken in the home (SP1) accounted for 2.2% of the variance

(R^2=0.022, $p< 0.001$) in science achievement. For Science 2 at Time 2, Spanish spoken in the home (SP2) accounted for 3.2% of the variance (R^2=0.032, $p< 0.001$) in science achievement. For Science 3 at Time 3, Spanish spoken in the home (SP3) accounted for 4.2% of the variance (R^2=0.042, $p< 0.001$) in science achievement. As for Science 4 at Time 4, Spanish spoken in the home (SP4) accounted for 10.3% of the variance (R^2=0.103, $p< 0.001$) in science achievement.

As for the executive function measure of numbers reversed (NRS2) at Time 2 on Math 2 accounts for 0.5% of the variance ((R^2=.005, $p< 0.001$) in math achievement. As for the executive function measure of numbers reversed at Time 1, Time 3, and Time 4, the numbers reversed task is not significant for mathematics achievement. As for science achievement, executive function measure of numbers reversed (NRS2) at Time 2 on Math 2 accounts for 0.8% of the variance ((R^2=.008, $p< 0.001$) in science achievement. As for the executive function measure of numbers reversed at Time 1, Time 3, and Time 4, the numbers reversed task is not significant for science achievement. For the card sort task, measuring working memory, Math 1 at Time 1, the card sort task (CS1) accounted for 4.0% of the variance (R^2=.04, $p< 0.001$) in mathematics achievement. Also yielding the same result, for the card sort task, Math 2 at Time 2, the card sort task (CS2) accounted for 4.0% of the variance (R^2=.04, $p< 0.001$) in mathematics achievement. At Time 3, the card sort task (CS3) accounted for 1.7% of the variance (R^2=.017, $p< 0.001$) in mathematics achievement (Math 3). At Time 4, the card sort task (CS4) accounted for 2.0% of the variance (R^2=.02, $p< 0.001$) in mathematics achievement (Math 4). As for science achievement, the card sort task (CS1 and CS2) accounts for 5.0% of the variance

(R^2=.05, p< 0.001) at times Time 1 and Time 2. At Time 3 and Time 4, the card sort task (CS3 and CS4), account for 4.0% (R^2=.05, p< 0.001) of the variance in science achievement. SP1-SP4 corresponds to Spanish spoken in the home, which is a dichotomous variable, NRS1-NRS4 is the numbers-reversed task at the four time points, the CS1-CS4 corresponds to the card sort task along the four time points. A value less than 0.001 is considered significant. The r-square value is multiplied by 100 to obtain the percentage of the variance in the achievement for mathematics and science.

Table 19
Various Covariates on Math and Science Along Four Time Points

Math 1 ON	Estimate	Std. Error	*p*-value	R2
SP1	0.05	0.01	0.000	0.003
NRS1	-0.02	0.04	0.643	
CS1	0.21	0.02	0.000	0.04
Math 2 ON				
SP2	0.045	0.01	0.000	0.002
NRS2	-0.07	0.02	0.000	0.005
CS2	0.192	0.01	0.000	0.04
Math 3 ON				
SP3	0.037	0.01	0.001	0.001
NRS3	0	0.019	0.951	
CS3	0.13	0.01	0.000	0.017
Math 4 ON				
SP4	0.05	4.24	0.000	0.003
NRS4	0.013	0.02	0.587	
CS4	0.123	0.012	0.000	0.02

Table 19 Continued

SCI 1 ON				
SP1	0.147	0.038	0.000	0.022
NRS1	-0.026	0.057	0.643	
CS1	0.072	0.005	0.000	0.005
SCI 2 ON				
SP2	0.179	0.047	0.000	0.032
NRS2	-0.091	0.024	0.000	0.008
CS2	0.068	0.004	0.000	0.005
SCI 3 ON				
SP3	0.206	0.061	0.001	0.042
NRS3	-0.002	0.026	0.951	
CS3	0.062	0.005	0.000	0.004
SCI 4 ON				
SP4	0.321	0.075	0.000	0.103
NRS4	0.019	0.035	0.587	
CS4	0.066	0.006	0.000	0.004

Results Based on Research Questions

- Is the initial level (*intercept* a.k.a. the "mean") for math and science significantly different for students in the non-Spanish speaking home environment as compared to students who speak Spanish dominantly in the home? (See Table 20). Are the intercepts statistically different in the children in the Spanish speaking versus the non-Spanish speaking under the unconditional model? Statistical significance of the intercept in the *unconditional math model* indicated the average starting value of math in this sample was significantly different than zero (2.89, $p < 0.001$).

Adding covariates to the model resulted in the conditional model being statistically insignificant. Significance of the intercept in the *unconditional science model* indicated the average starting value of science in this sample was significantly different than zero (3.35, $p < 0.001$). Adding the covariates had no effect on the science model did not change significance of the model. Also, the intercept in the *conditional science model* indicated the average starting value of science in this sample was significantly different than zero (3.3, $p < 0.001$).

- Is there significant *variance* of the intercept (a.k.a. the "mean") in math and science? If so, is it higher for students in the non-Spanish speaking home environment as compared to students who speak Spanish dominantly in the home? (See Table 20) Using the full PPLGCM with covariates, a finding of statistically significant variance in the *intercepts* based on the outcome of this study indicates that different students have statistically significant variance in the intercepts for math and science is interpreted as "some kids have score below the mean *initially*, some higher and some right at the mean" – that is, there is significant variation around the mean. The statistically significant difference between the intercepts and their random variation (0.01, $p < 0.001$) in the *unconditional math model* indicated that children differed with respect to their baseline math scores. The significant difference between intercepts and variance (0.01, $p < 0.001$) in *the unconditional science model* indicated that respondents differed with respect to their baseline science scores. The significant variance

(0.01, $p < 0.001$) in the *conditional science model* indicated that respondents differed with respect to their baseline science scores.

- Is the slope (rate of change) in math and science significantly different for students in the non-Spanish speaking home environment as compared to students who speak Spanish dominantly in the home? In the conditional math model, the mean of the latent slope factor indicated significant average growth (i.e. increase in scale score points) in math over time (6.5, $p < 0.001$). In the unconditional science model, the mean of the latent slope factor indicated significant average growth (i.e. increase in scale score points) in science over time (2.44, $p < 0.001$). In the conditional science model, the mean of the latent slope factor indicated average growth in science over time (2.6, $p < 0.001$). Using the full PPLGCM with covariates, a finding of statistically significant variance in the *slopes* for math and science is interpreted as "some kids have higher rate of change compared with others having a lower rate of change, and some kids maintain a "level" of flat rate of change (i.e. the children exhibit significantly different rates of change over time).
- Is the variance in the slope (variation around the rate of change) in math and science significantly different for students in the non-Spanish speaking home environment as compared to students who speak Spanish dominantly in the home? In the conditional math model, significant variance of the slope (0.68, $p < 0.001$) was observed indicating that not all students followed the same trajectory over the four time points. In the unconditional science model, the significant

variance of the slope (0.04, $p < 0.001$) indicated significant variability in average growth rates over time; not all students followed the same trajectory over the four time points. In the conditional science model, the significant variance of the slope (0.06, $p < 0.001$) indicated the variability in the average growth over time; not all students followed the same trajectory over the four time points.

Table 20
Research Questions 1-4

Model	Means	Estimate	S. E.	*p*-value	Effect size on a z-scale	Variance
Math Model without Spanish Home	Intercept	2.89	0.103	0.000	8.83	0.01
Math Model with Spanish Home	Intercept	0.49	0.44	0.266		0.19
Science Model without Spanish Home	Intercept	3.35	0.08	0.000		0.01
Science Model with Spanish Home	Intercept	3.30	0.070	0.000	0.62	0.01
Math Model without Spanish Home	Slope	9.4	3.58	0.009		12.80
Math Model with Spanish Home	Slope	6.5	0.82	0.000	1.32	0.68
Science Model without Spanish Home	Slope	2.44	0.21	0.000		0.04
Science Model with Spanish Home	Slope	2.6	0.24	0.000	-0.72	0.06

- Is there a significant correlation between the intercept and slope for math scores? If so, is it positive or negative? In the conditional mathematics model, the correlation between initial status and growth was significant and positive (r=0.119, SE=0.09, p=0.040). The small correlation between initial status and growth rate was only modest. The correlation between intercept and slope of the unconditional mathematics model was not statistically significant (r=0.753, SE=0.506, p=0.137). Using the full PPLGCM with covariates, a finding of statistically significant positive correlation between intercepts and slopes for math is interpreted as "on average students' level of math achievement increases over time". Student level of achievement over time decreased with the addition of the three covariates in the conditional model.

- Is there a significant correlation between the intercept and slope for science scores? If so, is it positive or negative? The correlations between intercept and slope of the conditional science model (r=0.032, SE=0.086, p=0.709) and unconditional science model (r=0.055, SE=0.088, p=0.530) are non-statistically significant. Using the full PPLGCM with covariates, a finding of statistically significant positive correlation between intercepts and slopes for science is interpreted as "on average students' level of science achievement increases over time". Results revealed no statistical significance for both models. On average, student level of science achievement decreased over time.
- Is there a significant correlation between the intercepts and slopes for math and science scores in the multivariate PPLGCM? If so, is it positive or negative? A significant positive correlation between intercepts in mathematics and science emerged in the conditional model (r=0.623, SE=0.029, p=0.000). A significant positive correlation between intercepts in mathematics and science also emerged in the unconditional model (r=0.66, SE=0.032, p=0.000). A significant positive correlation between slopes in mathematics and science occurs in the conditional

model (*r*=0.579, SE=0.081 *p*=0.000). A significant positive correlation occurred between slopes in mathematics and science occurs in the unconditional model (*r*=0.75, SE=0.029, *p*=0.010).

Table 21
Research Questions 5-7

Correlations	Model	Estimate	S. E	*p*-value	Effect size	Variance
$Intercept^{M}$ WITH $Slope^{M}$	Covariates	0.19	0.09	0.040		0.01
$Intercept^{M}$ WITH $Slope^{M}$	No Covariates	0.75	0.51	0.137	1.89	0.26
$Intercept^{S}$ WITH $Slope^{S}$	Covariates	0.03	0.09	0.709		0.01
$Intercept^{S}$ WITH $Slope^{S}$	No Covariates	0.06	0.09	0.530	0.26	0.01
$Intercept^{S}$ WITH $Intercept^{M}$	Covariates	0.62	0.03	0.000		0.00
$Intercept^{S}$ WITH $Intercept^{M}$	No Covariates	0.66	0.03	0.000	1.21	0.00
$Slope^{S}$ WITH $Slope^{M}$	Covariates	0.58	0.08	0.000		0.01
$Slope^{S}$ WITH $Slope^{M}$	No Covariates	0.75	0.29	0.010	0.92	0.08

- Does executive functioning serve as a significant covariate for math and science score performance in the PPLGCM? If so, how much variance in math or science achievement does it explain? As for the first measure of executive function, the numbers reverse (measure of cognitive flexibility) task, the only significant covariate was observed on the second assessment at Time 2 (*r*=-0.08, SE=0.02, *p*=0.000). Inclusion of the covariate only explained 0.6% of the variance in mathematics and science achievement at Time 2. As for the second measure of executive function, the card sort game (measure of working memory) was observed as statistically significant at Time 1 (r=0.21, SE=0.02, p=0.000), Time 2 (r=0.19, SE=0.01, p=0.000), Time 3 (r=0.13, SE=0.01, p=0.000) and Time 4 (r=0.12, SE=0.01, p=0.000) for both mathematics and science during the four time points. There is downward trend in how much variation the card sort function is explaining mathematics and science throughout the four time points. The numbers reverse NRS2 executive function covariate explains 0.6% of the variance in mathematics and science at Time 2. The card sort (CS1) executive function covariate explains 4.4% of the variance in mathematics and science at Time 1. The card sort (CS2) executive function covariate explains 3.6% of the variance in mathematics and science at Time 2. The card sort (CS3) executive function covariate explains 1.7% of the variance in mathematics and science at Time 3. The card sort (CS4) executive function covariate explains 1.4% of the variance in mathematics and science at Time 4. There is a downward trend as along the four time points suggesting the card sort game explains less of the variation as the children get older.

Table 22

Research Question 8

Interval	Covariate	Observed Variable	Estimate	S. E.	Variance	*p*-value
Time 1	NRS1	Math 1, Science 1	-0.02	0.04	0.000	0.643
Time 2	NRS2	Math 2, Science 2	-0.08	0.02	0.000	0.000
Time 3	NRS3	Math 3, Science 3	-0.00	0.02	0.000	0.951
Time 4	NRS4	Math 4, Science 4	0.01	0.02	0.000	0.587
Time 1	CS1	Math 1, Science 1	0.21	0.02	0.000	0.000
Time 2	CS2	Math 2, Science 2	0.19	0.01	0.000	0.000
Time 3	CS3	Math 3, Science 3	0.13	0.01	0.000	0.000
Time 4	CS4	Math 4, Science 4	0.12	0.01	0.000	0.000
Time 1	Sp1	Math 1, Science 1	0.05	0.01	0.000	0.000
Time 2	Sp2	Math 2, Science 2	0.05	0.01	0.000	0.000
Time 3	Sp3	Math 3, Science 3	0.04	0.01	0.000	0.001
Time 4	Sp4	Math 4, Science 4	0.05	0.01	0.000	0.000

- Does a PPLGCM without executive functioning covariates fit statistically better or worse than the model with covariates?

Table 23

Goodness-of-Fit Indicators for Five Models

Model	$\chi2$	df	*p* value	$\chi2$/df	RMSEA	90% CI	CFI	SRMR
Conditional Multivariate Model	2247.496	112	0.0000	20.067	0.083	0.080, 0.086	0.905	0.206
Unconditional Multivariate Model	1551.519	12	0.0000	129.293	0.215	0.206, 0.224	0.918	0.053
	$\Delta\ \chi^2$ = 695.977	Δdf = 100						

Note. RMSEA=root-mean-square of approximation; CFI=comparative fit index;
SRMR=standardized root mean square residual
***p < .001

To answer this question, a chi-square difference test between the two models is utilized. There are two models, conditional model with all covariates ($\chi2$=2247.5, *df*=112,

$p < 0.001$) and the unconditional model with no covariates ($\chi2=1551.5$, *df*=12, $p < 0.001$). The model with covariates has more degrees of freedom with the addition of the covariates. The model without the covariates fits statistically better than the model with covariates based on its chi-square value, which is almost half of the size of the model with covariates. If you look in a chi-square table and look for significance by taking the difference in degrees of freedom, this difference exceeds the critical number in the table. Therefore, the model with fewer covariates fits statistically better than the model with covariates in a structural equation modeling approach. The conditional multivariate model (with covariates, see Table 15) has a value of 0.905 and the unconditional multivariate model (no covariates) has a value of 0.918. Model fit indices indicate adequate fit of the unconditional linear model ($\chi2=1551.52$, *df*=12, $p < 0.001$; CFI=0.918, RMSEA=0.215; 90% CI [0.206, 0.22], SRMR=0.05). For the present study, the unconditional multivariate model has an SRMR value of 0.05 indicating a good fit for the model without covariates. Model fit indices indicate good fit of the conditional linear model ($\chi2=2247.5$, *df*=112, $p < 0.001$; CFI=0.91, RMSEA=0.08; 90% CI [0.08, 0.09], SRMR=0.21). For purposes of the present study (see Table 15) the multivariate conditional model, also known as the PPLGCM has an RMSEA value of 0.08 indicates an acceptable fit because a value closer to zero indicates the best fit. The models perform significantly different in terms of the CFI and RMSEA. This is important because although the chi-square difference was statistically significant, both models may fit pretty much the same in terms of the CFI and RMSEA values.

Presentation of results have included 1) Means, standard deviations, and sample sizes by covariates at each time point; 2) a summary of statistical estimates, standard errors, p-values, and model fit statistics for all parts of the PPLGM without covariates and 3) a summary of statistical estimates, standard errors, p-values, and model fit statistics for all parts of the PPLGM with covariates. We anticipate significant effects for means and variance components modeling in the PPLGM with and without covariates and data in the present study supports this.

Summary

Secondary data analysis was performed using data from the Early Childhood Longitudinal study. This analysis was designed to use parallel process latent growth curve modeling (LGCM) was conducted using a stepwise approach (Bollen & Curran, 2006). First, univariate refers to the analysis of mathematics and science achievement scores in isolation. Second, unconditional multivariate analyses contain information about the joint associations between growth factors of mathematics and science. Finally, conditional multivariate analyses incorporated the time-variant and time-invariant covariates as predictors of science and mathematics scores at each year. The full model for each type of analysis was presented in Figures 13-15. Measures assessed within the same occasion are considered to covary.

First, two unconditional (no covariates added) LGCM were specified to assess the change over time in the two outcomes of interest: Mathematics and Science achievement. Linear and nonlinear models were tested to find the best fit for the observed data patterns. Modification indices and specifications were also examined and adjusted to improve model fit. The fit of models was assessed using multiple fit indices that are sensitive to model misspecification in latent growth models. These indices include the RMSEA, CFI, SRMR. Demographic data summaries and descriptive analyses were performed in SPSS 24.0 and longitudinal growth curve models were conducted in MPlus 7.0 (Muthén & Muthén 1998-2014). Misssing data in all models were managed with the full-information maximum likelihood estimation utilized by MPlus. Unconditional LGCM included two latent (represented by ovals in Figures 14-19) variables. The latent growth factors were assumed to covary. To test whether the growth parameters of one curve was associated with the growth parameters of the other, the best-fitting models were then combined into one unconditional parallel process model. Covariates were added as a final step. The full information maximum likelihood estimator was utilized to account for missing data. To account for non-normality and skewness of variables, all models used maximum likelihood estimation with robust standard errors (Muthén & Muthén 1998-2014). Next, the research questions were again presented followed by tables answering each of the research questions supported by analytic results using ECLS-K data.

V. Discussion

Introduction

To build and sustain an effective STEM workforce, the recruitment, development and retention of unrepresented groups is essential. As STEM related occupations are projected to rise in the next decade, the need develop and implement systemic changes that provide access to quality educational experiences for students of color is critical. Given the occurring and projected demographic changes occurring in the nation, educational policy and practice must focus on establishing a culture that actively works to diversify the STEM workforce. As national discourses continue to demand educational institutions to produce creative and innovative minds, responses to this demand must be met with equally creative and innovative approaches. Establishing intentional and committed educational reform efforts aimed at raising the academic achievement of Hispanic youth in math and science can advance the goal of meeting the STEM workforce needs. This effort to shift the cultural landscape in STEM education begins investing deeply in and introducing Hispanic youth to STEM related professions and opportunities and providing rich learning experiences in early childhood education. These attempts to increase the number of the Hispanic population in STEM must also consider the racial and ethnic diversity of this particular population. Many Hispanic students are English language learners and as such, educational and research initiatives must acknowledge, value, and leverage racial and ethnic diversity of young people. The drive to answer a call to the shortage of a STEM workforce is of grave importance and it is imperative to establish a strong STEM pathway throughout the P-20 pipeline. The present study seeks to take a proactive stance to capitalize on students' language skills.

Research findings suggest that Hispanic students as group, are provided fewer opportunities in schools to acquire high-order skills in mathematics and science compared to their White counterparts (Clark, 1999, Jensen, 2007; Lindholm-Leary & Borsato, 2001; Lindholm-Leary & Borsato, 2005; Marian, Shook, & Schroeder, 2013; Strutchens & Silver, 2000; Tochon, 2009). One of the opportunities not given to Hispanic students as a whole is to develop their mathematics and science skills early on by building on their language development. Hispanic students are not given the opportunity to develop high-order thinking skills in mathematics and science.

Although much of the research in STEM education has collectively aimed to address issues related to STEM achievement and success, this research endeavor marks the initial attempt to utilize a longitudinal parallel process growth modeling to assess students in mathematics and science achievement and critically examine the language and executive function variables.

Existing literature focuses on distinct elements of STEM education but has not focused on assessing students in mathematics and science longitudinally through parallel process growth curve modeling, particularly with measures of executive functioning and Spanish in the home. This study is the first to follow children longitudinally in their early childhood trajectory examining the change over time in their mathematics and science achievement and whether Spanish and executive function significantly affect their achievement.

Little is known about language minorities' executive functioning and their achievement in science and mathematics. There has been some research that distinguish groups who purse

STEM fields compared to groups who do not pursue STEM fields by individual characteristics such as students' math and science attitudes, self-efficacy in math and science, gender, race/ethnicity, and structural characteristics including socioeconomic status, immigrant generation status, prior achievement in math and science, tracking, course taking patterns, and extracurricular involvement. However, there has been no research that differentiates groups by language proficiency (i.e. English speakers and non-native English speakers), specifically in the early grades. The gap in the literature informs the current proposed study. The proposed research study will broaden understanding, inform educational research, and impact classroom practice. The contents and the intended outcomes of this study will serve to inform decisions that affect educational policy, research, and practice and facilitates systemic change within our American public education system.

Review of Study Logic and Design

During the fall and spring, students in kindergarten, first and second grade took the mathematics and science assessment of the ECLS-K: 2011. In both the fall and the spring, trained and certified assessors assess students' executive function skills as well as ask the parents whether Spanish is spoken at the home or not. The battery of assessments was designed to be administered within about 60 minutes per child. Analyses were based on vertically equated scaled scores that are comparable from grade to grade and were expressed in Rasch units called RIT scores. The 20-item science assessment included questions about physical science, life sciences, environmental sciences, and scientific inquiry. The mathematics assessment measure conceptual knowledge, procedural knowledge, and problem solving through items related to number properties, operations, geometry, and spatial sense. Two measures of executive function were included in the kindergarten, first grade, and second-grade assessment battery. The *dimensional change card sort* (DCCS) and the (Zelaso, 2006; Zelaso et al 2013) and the Numbers Reversed subtest of the *Woodcock-Johnson III Tests of Cognitive Abilities* (Woodcock, McGrew, and Mather 2001), assessing working memory.

This study specifically investigates two variables of EF-working memory (WM) and cognitive shifting also known as cognitive flexibility by utilizing a parallel process longitudinal growth modeling to simultaneously study mathematics and science trajectories over time in a large cohort of students and to examine several critical issues related to potential changes in achievement over time. Using the PPLGM, the outcomes are threefold based on three types of analyses each serving a different purpose: 1) To examine the growth of mathematics or science in isolation employing a univariate analysis model within the PPLGM, 2) To reveal the joint associations between growth factors of mathematics and science employing an unconditional multivariate analysis and 3) To examine the effect of time-varying covariates as predictors of mathematics achievement scores at each year by employing a conditional multivariate analysis. Structural equation modeling (SEM) serves as the analytic framework for conducting our analyses. This study utilizes variables from the Early Childhood Longitudinal Study Cohort 2011. The Early Childhood Longitudinal Study is sponsored by the National Center for Education Statistics (NCES), within the U.S. Department of Education's Institute of Education Sciences, to provide detailed information on the school achievement and experiences of students throughout their elementary school years (McCarroll, Flanagan, & Potter, 2016). The students participating in the ECLS-K: 2011 are assessed longitudinally from kindergarten (the 2010-2011 school year) through the spring of 2016, when most are expected to be in fifth grade.

Review of Literature Findings

The study is grounded in a post-positivist framework that epistemological framework guiding the study and that emerged as a reaction of educational researchers to the limitations of positivism as a paradigm (Abdul Hameed, Sanaullah, & Asif Ali 2017). Educational researchers contended with the limitations of positivism for social sciences' research and combined positivism with interpretivism to form a new paradigm named post-positivism (Petter & Gallivan, 2004; Deluca, Gallivan, & Kock, 2008). The post-positivist critical realist firmly believes that "the goal of science is to hold steadfastly to the goal of getting it right about reality, even though we can never achieve that goal" (Trochim, 2006). Since most observation and measurement is fallible and all theory is revisable, the post-positivist emphasizes the importance of multiple measures and observations, each of which may possess different types of error (Trochim, 2006).

Revisiting the literature on executive function and achievement, Blair and Razza's (2007) study with preschool children in Head Start found that inhibitory control skills (a measure of executive function) was a significant predictor of mathematical skills rather than reading skills. The present study confirms that the two measures of executive function are correlated with mathematics and science achievement. As supported by the correlations table, the measure of executive function-the numbers reverse task measuring cognitive flexibility is highly correlated to mathematics achievement (r's> 0.5) in all four time points. Also consistent with the literature, the measure of the measure of executive function-the card sort task measuring working memory is moderately correlated to mathematics achievement (r's> 0.3) in all four time points. The findings of the present study are triangulated with similar findings in other studies where bilingual children exhibit increased benefits of executive functioning skills (Bialystok, 2007; Bialystok & Martin, 2004; Bialystok & Viswanathan, 2009; Carlson & Meltzoff, 2008), which correlate with math performance (Blair & Razza, 2007; Bull, Espy, Wiebe, 2008; Marian, Shook, Schroeder, 2013; Mazzoco & Kover, 2007; McClelland et al., 2007; Passolunghi & Siegel, 2001). Researchers Bull and Scerif (2001) assessed third graders with executive functioning tasks (e.g. card sorting task) and with a mathematics test of addition and subtraction; and through multiple linear regression analyses, they found that executive functioning reliably predicted mathematics performance. As for as science achievement, and consistent with the literature, the measure of executive function-the numbers reverse task measuring cognitive flexibility is moderate to highly correlated to science achievement (*r*'s> 0.4) in all four time points. Consistent with the science literature, the measure of executive function-the card sort task measuring working memory is highly correlated to science achievement (*r*'s> 0.3) in all four time points. As for Spanish of the home and mathematics achievement, Spanish of the home has a weak correlation to mathematics achievement (*r*'s> 0.2) in all four time points. For science achievement, Spanish of the home throughout the four time points has a moderate correlation (*r*'s> 0.4) with science achievement.

Discussion of the Results

A parallel process latent growth curve (LCCM) was conducted using a stepwise approach (Bollen & Curran, 2006). First, two unconditional (no covariates added) LGCM were specified to assess the change over time in the two outcomes of interest: Mathematics and Science achievement. Linear and nonlinear models were tested to find the best fit for the observed data patterns. Modification indices were also examined to improve model fit. Unconditional LGCM include two latent (represented by ovals in Figures 14-19). The latent growth factors were assumed to covary. To test whether the growth parameters of one curve was associated with the growth parameters of the other, the best-fitting models were then combined into one unconditional parallel process model. Covariates were added as a final step. The full information maximum likelihood estimator was utilized to account for missing data. To account for nonnormality and skewness of variables, all models used maximum likelihood estimation with robust standard errors (Muthén & Muthén 1998-2014). Using the full PPLGCM with covariates, a finding of statistically significant variance in the intercepts (as in the case of the present study) means that different students have statistically significant variation around the intercept (mean score) for math and science. This finding is interpreted as “some kids have scores below the mean *initially*, some higher and some right at the mean” – that is, there is significant variation around the mean. Using the full PPLGCM with covariates, a finding of statistically significant variance in the slopes for math and science (rates of change) science is interpreted as “some kids have higher rate of change compared with others having a lower rate of change, and some kids maintain a “level” of flat rate of change.

For Math 1 at Time 1, Spanish spoken in the home (SP1) accounted for 0.3% of the variance (R^2=.003, $p< 0.001$) in mathematics achievement. For Math 2 at Time 2, Spanish spoken in the home (SP2) accounted for 0.3% of the variance (R^2=.003, $p< 0.001$) in mathematics achievement. For Math 3 at Time 3, Spanish spoken in the home (SP3) accounted for 0.2% of the variance (R^2=.002, $p< 0.001$) in mathematics achievement. For Math 3 at Time 3, Spanish spoken in the home (SP3) accounted for 0.1% of the variance (R^2=.001, $p< 0.001$) in mathematics achievement. For Math 4 at Time 4, Spanish spoken in the home (SP4) accounted for 0.3% of the variance (R^2=.003, p< 0.001) in mathematics achievement. As for Science 1 at Time 1, Spanish spoken in the home (SP1) accounted for 2.2% of the variance (R^2=0.022, p< 0.001) in science achievement. For Science 2 at Time 2, Spanish spoken in the home (SP2) accounted for 3.2% of the variance (R^2=0.032, p< 0.001) in science achievement. For Science 3 at Time 3, Spanish spoken in the home (SP3) accounted for 4.2% of the variance (R^2=0.042, p< 0.001) in science achievement. As for Science 4 at Time 4, Spanish spoken in the home (SP4) accounted for 10.3% of the variance (R^2=0.103, p< 0.001) in science achievement.

As for the executive function measure of numbers reversed (NRS2) at Time 2 on Math 2 accounts for 0.5% of the variance ((R^2=.005, p< 0.001) in math achievement. As for the executive function measure of numbers reversed at Time 1, Time 3, and Time 4, the numbers reversed task is not significant for mathematics achievement. As for science achievement, executive function measure of numbers reversed (NRS2) at Time 2 on Math 2 accounts for 0.8% of the variance ((R^2=.008, p< 0.001) in science achievement. As for the executive function measure of numbers reversed at Time 1, Time 3, and Time 4, the numbers reversed task is not significant for science achievement.

For the card sort task, measuring working memory, Math 1 at Time 1, the card sort task (CS1) accounted for 4.0% of the variance (R^2=.04, p< 0.001) in mathematics achievement. Also yielding the same result, for the card sort task, Math 2 at Time 2, the card sort task (CS2) accounted for 4.0% of the variance (R^2=.04, p< 0.001) in mathematics achievement. At Time 3, the card sort task (CS3) accounted for 1.7% of the variance (R^2=.017, p< 0.001) in mathematics achievement (Math 3). At Time 4, the card sort task (CS4) accounted for 2.0% of the variance

(R^2=.02, p< 0.001) in mathematics achievement (Math 4). As for science achievement, the card sort task (CS1 and CS2) accounts for 5.0% of the variance (R^2=.05, p< 0.001) at times Time 1 and Time 2. At Time 3 and Time 4, the card sort task (CS3 and CS4), account for 4.0% (R^2=.05, p< 0.001) of the variance in science achievement.

Using the full PPLGCM with covariates, a finding of statistically significant positive correlation between intercepts and slopes for math is interpreted as "on average students' level of math achievement increases over time". For purposes of this study, one of the objectives was to examine whether math achievement increases over time in the same pattern for non-Spanish speaking students in the home compared to those students that speak Spanish dominantly in the home. Using the full PPLGCM w/covariates, a finding of statistically significant positive correlation between intercepts and slopes for science is interpreted as "on average students' level of science achievement increases over time". *For purposes of this study, one of the objectives was to examine whether science achievement increases over time in the same pattern for non-Spanish speaking students in the home compared to those students that speak Spanish dominantly in the home.

The significant mean of the intercept in the *unconditional math model* indicated the average starting value of math in this sample was significantly different than zero (2.89, $p <$ 0.001). Adding covariates to the model made the now conditional model insignificant. Adding covariates to the model made the model more complex and strained. The significant mean of the intercept in the *unconditional science model* indicated the average starting value of science in this sample was significantly different than zero (3.35, $p < 0.001$). Adding the covariates had no effect on the science model did not change significance of the model. The significant mean of the intercept in the *conditional science model* indicated the average starting value of science in this sample was significantly different than zero (3.3, $p < 0.001$). The significant variance (0.01, $p < 0.001$) in the *unconditional math model* indicated that respondents differed with respect to their baseline math scores. The significant variance (0.01, $p < 0.001$) in *the unconditional science model* indicated that respondents differed with respect to their baseline science scores. The significant variance (0.01, p < 0.001) in the *conditional science model* indicated that respondents differed with respect to their baseline science scores. In the conditional math model, the mean of the latent slope factor indicated average growth in math over time (6.5, $p < 0.001$). In the unconditional science model, the mean of the latent slope factor indicated average growth in science over time (2.44, $p < 0.001$). In the conditional science model, the mean of the latent slope factor indicated average growth in science over time (2.6, $p < 0.001$). In the conditional math model, the significant variance of the slope (0.68, $p < 0.001$) indicated the variability in the average growth over time; not all students followed the same trajectory over the four time points. In the unconditional science model, the significant variance of the slope (0.04, $p < 0.001$) indicated the variability in the average growth over time; not all students followed the same trajectory over the four time points. In the conditional science model, the significant variance of the slope (0.06, $p < 0.001$) indicated the variability in the average growth over time; not all students followed the same trajectory over the four time points. In the conditional mathematics model, the correlation between initial status and growth was significant and positive (r=0.119, SE=0.09, p=0.040). The small correlation between initial status and growth rate was only modest. The correlations between intercept and slope of the unconditional mathematics model are non-statistically significant (r=0.753, SE=0.506, p=0.137). Using the full PPLGCM with covariates, a finding of statistically significant positive correlation between intercepts and slopes for math is interpreted as “on average students’ level of math achievement increases over time”. Student level of achievement over time decreased with the addition of the three covariates in the conditional model. The correlations between intercept and slope of the conditional science model (r=0.032, SE=0.086, p=0.709) and unconditional science model (r=0.055, SE=0.088, p=0.530) are non-statistically significant. Both models were non-statistically significant therefore, on average, student level of science achievement decreased over time. If so, how much variance in math or science achievement does it explain? As for the first measure of executive function, the numbers reverse (measure of cognitive flexibility) task the only significant covariate is the second assessment at Time 2 (r=-0.08, SE=0.02, p=0.000). It is only explaining 8% of the variance in mathematics and science achievement at Time 2. As for the second measure of executive function, the card sort game (measure of working memory) is significant in Time 1 (r=0.21, SE=0.02, p=0.000), Time 2 (r=0.19, SE=0.01, p=0.000), Time 3 (r=0.13, SE=0.01, p=0.000) and Time 4 (r=0.12, SE=0.01, p=0.000) for both mathematics and science during the four time points. There is downward trend in how much variation the card sort function is explaining mathematics and science throughout the four time points.

Assumptions and Limitations

There are some important caveats regarding the results. First, the data are merely correlational, and therefore, only limited evidence as to the causal mechanisms underlying the observed changes over time.

Strength of Study

Study strengths include a representative sample and the heterogeneity of variables available in the ECLS-K: 2011 dataset. This study used an innovative methodology for studying parallel process longitudinal growth curve modeling of a longitudinal assessment. Parallel process LGCM is an innovative model for examining effects in growth curve models. First, they were used to simultaneously model the progression trajectories of mathematics and science achievements with and without the effects of covariates. PPLGCMs allow pairwise correlations of the latent intercepts and slopes of outcome variables to be included in the models using a structural equation modeling framework. This approach improves model fit. Second, this approach allows estimation of individual differences and changes over time. This study is the first of its kind to have utilized such a method to analyze achievement in mathematics and science over time with the influence of the language and executive function covariates.

Future Research

Suggested future research includes considering novel approaches to studying mathematics and science achievement utilizing novel methodologies. Future research should delve further into the field neuropsychology, specifically executive functioning and the effects of different types of executive function on mathematics and science achievement. Future research should also examine science and mathematics separately to understand all the underlying variables contributing to the variance in achievement.

Significance

The problem of co-occurring factors (covariates) provides a challenge for determining effective student-centered interventions given that changes in such co-occurring factors occurs over time and thus influences mathematics and science achievement scores. Here we introduce an analytic approach the enables researchers to perform longitudinal growth analysis in light of the influence of co-occurring conditions over time as related to students' mathematics and science achievement scores. Developing methods for mediation analysis of co-occurring variables for complex data longitudinal data as in the Early Childhood Longitudinal Study can help researchers make improved use of databases like the ECLS-K: 2011 to better understand mechanisms of change in educational achievement. My proposed research will call on policymakers, researchers, and practitioners to consider low levels of academic achievement in either reading or mathematics as very common among the general population of U.S. schoolchildren by the early grades. My research will provide additional empirical support for efforts to help children experiencing the early onset of learning difficulties. The results of the present study indicate that kindergarten children with working memory and cognitive flexibility deficits in early childhood will impact their achievement in mathematics and science throughout their schooling. The covariates added in the model add little to the explanatory power of both models. Large effect size between the model with covariates and the model without covariates help explain the differences between both models (see Table 20). Spanish spoken in the home accounted for very little of the variance in mathematics achievement during the four time points, 0.3%, 0.3%, 0.2% and 0.1%, respectively. As for science achievement, results indicated Spanish of the home accounted for 2.2% during Time 1, 3.2% during Time 2, 4.2% during Time 3, and 10.3% of the variance during Time 4 suggesting Spanish explains more variation in science achievement rather than mathematics. The numbers reversed task at Time 2 only accounts for 0.5% of the variance in mathematics achievement. At Time 1, Time 3, and Time 4, the numbers reversed task is not a significant covariate for mathematics achievement. As for science achievement, executive function measure of numbers reversed (NRS2) at Time 2 only accounts for 0.8% of the variance in science achievement. The numbers reversed task is not significant at Time 1, Time 3, and Time 4 and therefore does not contribute to the variance in science achievement during those time points. Though the numbers reverse task contributed little to the overall variance in achievement of mathematics and science, the card sort task accounts a greater percentage of the variance in mathematics and science throughout the four time points. During Time 1 and Time 2, the card sort task accounted for 4.0% of the variance in mathematics achievement. At Time 3, the card sort task accounted for 1.7% of the variance in mathematics achievement. At Time 4, the card sort task accounted for 2.0% of the variance in mathematics achievement. As for science achievement, the card sort task accounted for 5.0% of the variance at times Time 1 and Time 2. At Time 3 and Time 4, the card sort task account for 4.0% of the variance in science achievement. Combined, the effects of the covariates account for 4.3% of the variance in mathematics achievement at Time 1, 4.7% of the variance in mathematics achievement at Time 2, 1.8% of the variance in mathematics achievement at Time 3, and 2.3% of the variance in mathematics achievement at Time 4. As for science achievement, the combined effects of the covariates account for 2.7% of the variance in science achievement at Time 1, 4.5% of the variance in science achievement at Time 2, 4.6% of the variance in science achievement at Time 3, and 10.7% of the variance in science achievement at Time 4.

The present study also confirms that deficits in working memory (numbers reversed task) are predictive of learning difficulties in mathematics and science in early childhood. It is necessary to begin addressing executive function deficits early in the educational trajectory of students in order to achieve greater success in both mathematics and science in general. Also, Spanish in the home along the four time points had a weak correlation in mathematics achievement. As for science achievement, Spanish in the home had moderately high correlation with science throughout the four assessment time points. Supported by the aforementioned results, it is essential to capitalize and build upon the home language of students in order to increase achievement in science.

The results also empirically support interventions that target EF (executive function) as an important component of early childhood mathematics and science education. Given the critical importance of executive function skills for early school success and later schooling outcomes, the current findings have implications for assessment and educational practices. As reviewed previously, individual differences in EF measured in childhood not only predict academic outcomes, but also predict other important outcomes, including long-term physical and mental health (Zelazo et al., 2016). There is also evidence that children with better EF skills actually learn more and retain more information from a given amount of instruction and practice (Zelazo et al., 2016).

Some of those interventions utilized to improve executive functioning in early school include computerized training, computer and non-computer games, aerobic exercise, physical activity, music training, martial arts, and mindfulness practices (Diamond & Lee, 2011). Two studies have evaluated the efficacy of EF training through specific sets of adult-led, game-like activities with prechoolers. These activities include typical children's games involving inhibitory control, cognitive flexibility, and working memory, similar to Simon Says and Red Light, Green Light (Zelazo et al., 2016). An initial evaluation of six activities with 65 preschool children delivered over 16 brief (20-30 minutes) playgroup sessions by a trained adult indicated no main effects of the activities on child EF (Tominey & McClelland, 2011). A second evaluation of the activities with 276 children in 14 pre-K classrooms (Schmitt et al., 2015) were associated with small to moderate gains on measures of EF, the DCCS (*d*=.16) and the Head-Toes-Knees-Shoulders task (*d*=.32). In the aforementioned study, the strong effects of the intervention on gains in math (*d*=.44) of the subset of children (*N*=88) who were identified as ELL.

Other than the aforementioned practices, it is necessary to delve into programs designed as comprehensive curricula which can be utilized to supplement existing practice (Zelazo et al., 2016). An example of a comprehensive approach to the education of young children focusing on early academic learning through the mechanism of EF is the Tools of the Mind program (Bodrova & Leong, 2007). Tools of the Mind blends teacher-led scaffolding of a comprehensive curriculum of early literacy, mathematics, and science activities with child-directed activities and structured sociodramatic play.

Implications for Policy and Practice

These findings have significant implications for educators, school leaders, and principals since much of the focus of current research on EF and education is on the way in which EF contributes to academic learning. It is increasingly important to consider not only the ways in which improvements in EF may lead to improvements in academic ability, but also the extent to which improvements in EF can contribute to language acquisition. A continued research and policy focus is needed on the measurement of EF and on trajectories of language and EF development from early childhood through young adulthood. Continued work on the longitudinal assessment of EF beyond the ECLS-K: 2011 is needed. Overall research suggests that EF and language provide a foundation for learning and adaptation is wide range of circumstances, including school. EF skills needed to be successful in school such as attentive listening, keeping information in mind, thinking flexibly, and inhibiting impulses can be acquired in school settings leading to improved academic achievement.

Summary and Conclusions

Parallel process longitudinal growth curve modeling was used to examine the longitudinal conditional effect of three covariates on mathematics and science achievement. The parallel process LGCM jointly modeled the trajectories of the covariates and the outcomes over time in a structural equation modeling framework, allowing for assessment of relationships among the latent factors the covariates and the latent factors of outcome. Parallel processes (i.e., growth in mathematics and science) were modeled simultaneously over four time points of the Early Childhood Longitudinal Study of 2011. A stepwise approach was adopted to establish the presence of covariates. A series of longitudinal growth curve models were fit. The univariate models on covariate and outcome along time were modeled first to check the model fit. The model specifications were adjusted as needed to improve model fit. The fit of the models was assessed using multiple fit indices that are sensitive to model misspecification in latent growth models. These indices included the root mean square error of approximation (RMSEA), the comparative fix index (CFI), and the standardized root mean square residual (SRMS).

REFERENCES

Abedi, J., & Gándara, P. (2006). Performance of English language learners as a subgroup in large scale assessment: Interaction of research and policy. *Educational Measurement*: *Issues and Practice*, *25*(4), 36-46.

Abedi, J., & Lord, C. (2001). The language factor in mathematics tests. *Applied Measurement in Education*, *14*(3), 219-234.

Abdul Hammed, P., Sanaullah, A., & Asif Ali, S. (2017). Post-positivism: an effective paradigm for social and educational research. *International Journal of Art & Humanities*, *45*(45), 253-259.

Ackerman, D. J., & Tazi, Z. (2015). Enhancing young Hispanic dual language leaners' achievement exploring strategies and addressing challenges. *ETS Research Reports Series*, 2015(1).

Agostino, A., Johnson, J., & Pascual-Leone, J. (2010). Executive functions underlying Multiplicative reasoning: Problem type matters. *Journal of Experimental Child Pyschology*, *105*, 286-305.

Allan, N. P., Hume, L. E., Allan, D. M., Farrington, A. L., & Lonigan, C. J. (2014). Relations between inhibitory control and the development of academic skills in preschool and kindergarten: A meta-analysis. *Developmental Psychology*, *50*(10), 2368-2379.

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

Anderson, E., & Kim, D. (2006). *Increasing the success minority students in science and technology*. Washington, DC: American Council on Education.

Angelova, M., Gunawardena, D., & Volk, D. (2006). Peer teaching and learning: Co-constructing language in a dual language first grade. *Language and Education, 20*(3), 173-190.

Archibald, J., Roy, S., Harmel, S., & Jesney, K. (2006) A review of the literature on second language learning (2nd ed.). Calgary, Alberta, Canada: University of Calgary, Language Research Center.

Armstrong, P. W., & Rogers, J. D. (1997). Basic skills revisited: The effects of foreign language instruction on reading, math and language arts. *Learning Languages*, *2*(3), 20-31.

Aronin, S & Floyd, K. K. (2013). Using an iPad in inclusive preschool classrooms to introduce STEM concepts. *Teaching Exceptional Children*, 45(4), 34-39.

August, D., & Hakuta, K. (Eds). (1997). Improving schooling for language-minority children: A research agenda. Washington, D.C.: National Academy Press.

August, D. (2006). Demographic overview. IN D. August & T. Shanahan (Eds.), Report of the National Literacy Panel on Language Minority Youth and Children. Mahwah, NJ: Lawrence Erlbaum.

Baker, C. (2011). Foundations of bilingual education and bilingualism (5th ed.). Clevedon, England: Multilingual Matters.

Bassok, D., Latham, S., & Rorem, A. (2016). Is kindergarten the new first grade? *AERA Open*, *1*, 1-31.

Bialystok, E. (2001). *Bilingualism in Development: Language, Literacy, and Cognition*. New

York: Cambridge University Press.

Bialystok, E. (2007). Cognitive effects of bilingualism: How linguistic experience leads to cognitive change. *International Journal of Bilingual Education and Bilingualism*, *10*(3), 210-223.

Bialystok, E. (2008). Language processing in bilingual children. New York: Cambridge University Press.

Bialystok, E., & Barac, R. (2012). Emerging bilingualism: Dissociating advantages for metalinguistic awareness and executive control. *Cognition*, *122*, 67-73.

Bialystok, E., & Martin, M. M. (2004). Attention and inhibition in bilingual children: Evidence from the dimensional change card sort task. *Developmental Science*, *7*(3), 325-339.

Bialystok, E., & Viswanathan, M. (2009). Components of executive control with advantages for bilingual children in two cultures. *Cognition*, *112*, 494-500.

Blair, C. (2002). School readiness: Integrating cognition and emotion in a neurobiological conceptualization of children's functioning at school entry. *American Psychologist*, *57*(2), 111-127.

Blair, C., & Diamond, A. (2008). Biological processes in prevention and intervention: Promotion of self-regulation and the prevention of early school failure. Development and Psychopathology, 20(3): 899-911.

Blair, C., & Raver, C. C. (2015). School readiness and self-regulation: A developmental psychobiological approach. *Annual Review of Psychology*, *66*: 711-731.

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

Blozis, S. A., Harring, J. R., & Mels, G. (2008). On fitting nonlinear latent curve models to multivariate longitudinal data. *Strucutral Equation Modeling, A Multidisciplinary Journal 15*, 356-379.

Bodrova, E., and Leong, D. J (2007). Tools of the mind: The Vygotskian approach to early childhood education (2nd ed.). New York: Merrill/Prentice Hall.

Bollen, K. A., & Curran, P. J. (2006). *Latent curve models: A structural equation perspective* (Vol. 467). Hoboken, New Jersey: Wiley.

Bornstein, M. H., Hahn, C. S., Putnick, D. L., & Suwalksky, J. T. (2014). Stability of core language skill from early childhood to adolescence: A latent variable approach. *Child Development*, *85*, 1346-1356.

Bruce, B. C., Bruce, S., Conrad, R., & Huang, H. (1997). Collaboration in science education: University science students in the elementary school classroom. *Journal of Research in Science Teaching*, *34*, 69-88.

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

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

Buysse, V., Peisner-Feinberg, E., Paez, M., Hammer, C. S. & Knowles, M. (2014). Review: Effects of early education programs and practices on the development and learning of dual language learners: A review of the literature. *Early Childhood Research Quarterly,* *29*, 765-785.

Byrne, B. (2010). Structural equation modeling with MPlus: Basic concepts, applications, and programming. New York: Routledge.

Campbell, F. A., Pungello, E. P., Miller-Johnson, S., Burchinal, M. & Ramey, C. T. (2001). The development of cognitive and academic abilities: Growth curves from an early childhood educational experiment. Developmental Psychology, 37(2), 231-242. doi: 10.1037//0-121649.37.2.231

Carlson, S. M., & Meltzoff, A. N. (2008). Bilingual experience and executive functioning in young children. *Developmental Science*, *11*(2), 282-298.

Carlson, S. M., Zelazo, P. D., & Faja, S. (2013). Executive Function. *Oxford Handbook of Developmental Psychology*, *Vol. 1* (pp. 06-742).

Carson, S. M., Zelazo, P. D. & Faja, S. (2013). Executive function. *Oxford Handbook of Developmental Psychology*, *Vol. 1* (pp.706-742.

Castro, D., Espinosa, L. M., & Paez, M. (2011). Defining and measuring quality early childhood practices that promote dual language learners' development and learning. In M. Zaslow, I. Martinez-Beck, K. Tout, & T. Halle (Eds.), *Next steps in the measurement of quality in early childhood settings* (pp.257-280). Baltimore, MD: Brookes.

Center on the Developing Child at Harvard University. (2011). Building the brain's "air traffic control" system: How early experiences shape the development of executive function (Working Paper No. 11). Retrieved from http://www.developingchild.harvard.edu

Chapman, J. W. (1988). Learning disabled children's self-concepts. *Review of Educational Research, 58,* 347-371.

Chessloff, J. D. (2013). Why STEM education must start in early childhood. *Education Week*, 32(23), 27-32.

Chetty, R., Friedman, J. N., Hilger, N., Saez, E., Whitemore Schanzenbach, D., & Yagan, D. (2011). How does your kindergarten classroom affect your earnings? Evidence from project STAR. *Quartely Journal of Economics, 126*, 1593-1660.

Choi, J. Y, Jeon, S., & Lippard, C. (2018). Dual language learning, inhibitory control, and math achievement in Head Start and kindergarten. *Early Childhood Research Quarterly*, 42, 66-78.

Christian, D. (1996). Two-way immersion education: Students learning through two languages. *The Modern Language Journal*, *80*(1), 66-76.

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

Clarkson, P.C. (1992). Language and mathematics: A comparison of bilingual and monolingual students of mathematics. *Educational Studies in Mathematics*, *23*(4), 417-429.

Clarkson, P.C. (2007). Australian Vietnamese students learning mathematics: High ability bilinguals and their use of their languages. *Educational Studies in Mathematics*, *64*, 417-429.

Cobb, B., Vega, D., & Kronauge, C. (2006). Effects of an elementary dual language immersion school program on junior high achievement. *Middle Grades Research Journal*, *1*, 27-47.

Collier, V. P. & Thomas, W. P. (2004). The astounding effectiveness of dual language education for all. *NABE Journal of Research and Practice*, *2*(1), 1-20.

Committee on Underrepresented Groups and the Expansion of the Science and Engineering Workforce Pipeline, et al. (2011). *Expanding underrepresented minority participation: America's science and technology talent at the crossroads*. National Academies Press.

Curran, P. J., & Hussong, A.M. (2003). The use of latent trajectory models in psychopathology research. *Journal of Abnormal Psychology*, *112*, 526-544.

Curtain, H., & Dahlberg, C. A. (2004). *Languages and children: Making the match* (3rd ed.). Boston: Pearson.

Dang, M. (2015). Examining the math attitudes, math self-efficacy, and STEM outcomes among native and non-native English speakers using a latent class analysis. (Order No. 3733577). Available from ProQuest Dissertations & Theses Global. (1731045075). Retrieved from http://libproxy.txstate.edu/login?url=https://search-proquest-com.libproxy.txstate.edu/docview/1731045075?accountid=5683

Darling-Hammond L. (2009). The flat world and education: How America's commitment to equity will determine our future. New York: Teachers College Press.

Davis, J., & Bauman, K. (2013). *School enrollment in the United States: 2011*. Washington, DC: U.S. Census Bureau.

DeJarnette, N. K. (2012). America's children: providing early exposure to STEM (science, technology, engineering and math) initiatives. *Education*, 133(1), 77-84.

Dempster, F. N. (1992). The rise and fall of the inhibitory mechanism: Toward a unified theory of cognitive development and aging. *Developmental Review*, *12*, 45-75.

Dennis, M. (1991). Frontal lobe function in childhood and adolescence: A heuristic for assessing attention regulation, executive control, and in the intentional states important for social discourse. *Developmental Neuropsychology*, *7*, 327-358.

Department of Education. National Center for Education Statistics (2011). The Nation's Report Card: Science 2009 (Report No. NCES 2011–451). Washington, DC: Institute of Education Sciences, U.S. Department of Education.

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*, 135-168.

Diamond, A., Barnett, S. W., Thomas, J., & Munro, S. (2007). Preschool program improves cognitive control. *Science*, *318*(5855), 1387-1388.

Diamond, A., & Lee, K. (2011). Interventions show to aid executive function development in children 4 to 12 years old. *Science*, *333*(6045), 959-964.

Dorney, Z. (2005). *The psychology of the language learner*. Mahwah, NJ: Lawrence Erlbaum.

Drake, T. A. (2014). The effect of community linguistic isolation on language-minority student achievement in high school. *Educational Researcher*, *43*(7), 327-340. doi:10.3102/0013189X14547349

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P.,...Japel, C. (2007). School readiness and later achievement. *Developmental Psychology*, *43*, 1428-1446.

Duncan, T.E., Duncan, S. C., Strycker, L. A., Li, F., & Alpert, A. (1999). *An introduction to latent variable growth curve modeling. Concepts, issues, and applications*. Mahwah, NJ: Erlbaum.

Duschl, R. A, Schweingruber, H. A., & Shouse, A. W. (2007). What research says about K-8 science learning and teaching. *Principal, 87*(2), 16-22.

Else-Quest, N.M., Mineo, C. C., & Higgins, A. (2013). Math and science attitudes and achievement at the intersection of gender and ethnicity. Psychology of Women Quarterly, 37(3), 293-309. doi:10.3102/0013189X14547349

Emeagwali, N. S. (2010). National science board says U. S. lead in STEM slipping. *Techniques: Connecting Education & Careers, 85*(3), 10-11.

Eshach, H. (2003). Inquiry-events as a tool for changing teaching efficacy belief of kindergarten and elementary school teachers. *Journal of Science Education and Technology*, *12*(4), 495-501.

Eshach, H., & Fried, M. N. (2005). Should science be taught in early childhood? Journal of *Science Education and Technology*, *14*(3), 315-336.

Eslinger, P.J. (1996). Conceptualizing, describing, and measuring components of executive function: A Summary. In G. R. Lyon & N. A. Krasnegor (Eds.), *Attention, memory, and executive function*. Baltimore: Brookes.

Espinosa, L. M. (2007). English-language learners as they enter school. In R. Pianta, M. Cox, & K. Snow (Eds.), *School readiness and the transition to kindergarten in the era of accountability* (pp.175-196). Baltimore, MD: Paul H. Brookes.

Esposito, A. G., & Baker-Ward, L. (2013). Dual-language education for low-income children: Preliminary evidence of benefits for executive function. *Bilingual Research Journal*, *36*, 295-310.

Espy, K. (2004). Using developmental, cognitive, and neuroscience approaches to understand executive control in young children. *Developmental Neuropsychology*, *26*, 379-384.

Excelencia in Education. (2015). *The condition of Latinos in education: 2015 Factbook.* Washington, D. C. : *Excelencia* in Education.

Fan, X., Thompson, B. & Wang, L. (1999). Effects of sample size, estimation methods, and model specifications on structural equation modeling fit indexes. *Structural Equation Modeling*, 6, 56-83.

Fan, X. & Konold, T. R. (2009). Latent growth curve analysis in structural equation modeling: Concepts and implementations. In Teo, T. & Khine, M. S. (eds.), *Structural Equation Modeling: Concepts and Applications in Educational Research* (pp 29-58).

Farrell, M. P. (2011). Bilingual competence and students' achievement in physics and mathematics. *International Journal of Bilingual Education and Bilingualism*, *14*(3) 335-345.

Fitts, S. (2006). Reconstructing the status quo: Linguistic interaction in a dual language school. *Bilingual Research Journal, 30*(2), 337-365.

Flores, A. (2017). How the U.S. Hispanic population is changing. Pew Research Center, Washington, D.C. Retrieved from: http://www.pewresearch.org/fact-tank/2017/09/18/how-the-u-s-hispanic-population-is-changing/

Fredricks, D. E., & Warriner, D. S. (2016). "We speak English in here and English only!": Teacher and ELL youth perspectives on restrictive language education. *Bilingual Research Journal*, *39*(3-4), 309-323.

Freire, J. A. (2016). Nepantleras/os and their teachers in dual language education: Developing sociopolitical consciousness to contest language education policies. *Association of*

Mexican American Educators Journal, *10*(1), 36-52.

Fry, R. (2007). How far behind in math and reading are English language learners? Washington, DC: Pew Hispanic Research Center.

Gándara, P., Rumberger, R., Maxwell-Jolly, J. & Callahan, R. (2003). English learners in California schools: Unequal resources, unequal outcomes. *Education Policy Analysis Archives*, *11*(36). Retrieved from http://epaa.asu/edu/epaa/v11n36/

Garcia, E., Arias, M. B., Harris Murri, N. J., & Serna, C. (2010). Developing responsive teachers: A challenge for a demographic reality. *Journal of Teacher Education*, *61*(1-2), 132-142.

Garcia E. E., & Nañez. J. E. Sr. (2011). Bilingualism and cognition: Informing research, pedagogy, and policy (pp. 57-77). Washington, DC: American Pyschological Association.

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

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

Gathercole, S.E., Lamont, & Alloway, T. P. (2006). Working memory in the classroom. In S.J. Pickering (Ed.), *Working memory and education*, 219-240. Oxford, UK: Elsevier.

Gathercole, S. E., Tiffany, C., Briscoe, J., Thorn, A., & ALSPAC team (2005). Developmental Consequences of poor phonological short-term memory function in childhood: A Longitudinal study. *Journal of Child Psychology and Psychiatry*, *46*, 598-611.

Genesee, F. (1983). Bilingual education of majority-language children: The immersion experiments in review. *Applied Psycholinguistics*, *4*(1), 1-46.

Genesee, F., Lindholm-Leary, K., Saunders, W. M., & Christian, D. (Eds.) (2006). Educating English language learning: A synthesis of research evidence. New York, NY: Cambridge University Press.

Ginsburg, H. P., & Golbeck, S. L. (2004). Thoughts on the future of research on mathematics and science learning and education. Early Childhood Research Quaterly, 19(1), 190-200.

Gomez, L., Freeman, D., & Freeman, Y. (2005). Dual language education: A promising 50-50 model. Bilingual Research Journal, 29, 145-64.

Gonzalez-Barrera, A. & Lopez, M. H. (2013). Spanish is the most spoken non-English language in U. S. homes, even among non-Hispanics. Pew Research Center. Washington, D.C. Retrieved from: http://www.pewresearch.org/fact-tank/2013/08/13/spanish-is-the-most-spoken-non-english-language-in-u-s-homes-even-among-non-hispanics/

Gormely, W., Gayer, T., Phillips, D., & Dawson, B. (2005). The effects of universal pre-K on cognitive development. *Developmental Psychology*, *41*, 872-884.

Greene, J. P. (1998). A meta-analysis of the effectiveness of bilingual education. Claremont, CA: Tomas Rivera Policy Institute.

Grimm, K. J., Ram, N., & Estabrook, R. (2017). Growth modeling: Structural equation modeling and multilevel modeling approaches. New York, NY: Guilford Press.

Grigg, W., Lauko, M., & Brockway, D. (2006). The Nation’s Report Card: Science 2005 (NCES 2006-466). US. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Guglielmi, R. S. (2012). Math and science achievement in English language learners: Multivariate latent growth modeling of predictors, mediators, and moderators. *Journal of Educational Psychology*, *104*(3), 580-602.

Hadzigeorgiou, Y. (2002). A study of the development of the concept of mechanical stability in preschool children. *Research in Science Education*, *32*(3), 373-391.

Halle, T., Hair, E., Wandner, L., McNamara, M., & Chien, N. (2012). Predictors and outcomes of early versus later English proficiency among English language learners. Early Childhood Research Quaterly, 27(1), 1-20.

Harvey, H. A., & Miller, G. E. (2017). Executive function skills, early mathematics, and vocabulary in head start preschool children. *Early Education and Development*, *28*(3), 290-307.

Hans-Vaughn, D. L. (2005). A primer for using and understanding weights with national datasets. *The Journal of Experimental Education*, *73*(*3*), 221-248.

Heck, R.H., & Thomas, S.L. (2015). *An introduction to multilevel modeling techniques: MLM and SEM approaches using MPlus., 3rd. ed.* New York, NW, US: Routledge/Taylor & Francis Group.

Hernandez, D. (2006). Young Hispanic children in the U.S.: A demographic portrait based on Census 2000. Albany: University at Albany, State University of New York.

Horizon Research (2013). *Report of the 2012 national survey of science and mathematics education*. Chapel Hill, NC: NSSME.

Hughes, C. (2011). Changes and challenges in 20 years of research into the development of executive functions. *Infant and Child Development*, *20*(3), 251-271.

Institute for Education Sciences & National Center for Education Statistics (2017). Early Childhood Longitudinal Program (ECLS) Child Assessments. https://nces.ed.gov/ecls/assessments2011.asp

Institute of Medicine, National Academy of Sciences, and National Academy of Engineering (2007). *Rising above the gathering storm: Energizing and employing America for a brighter economic future*. Washington, DC: The National Academies Press.

IOM, NAS, & NAE (2007). Rising above the gathering storm: Energizing and Employing America for a brighter economic future. Washington, DC: The National Academies Press.

Isquith, P., Crawford, J., Espy, K., & Gioia, G. (2005). Assessment of executive function in preschool-aged children. *Mental Retardation & Development Disabilities Research Reviews*, *11*, 209-215.

Jacques, S., & Marcovitch, S. (2010). Development of executive function across the life span. In W. Overton (Ed.), *Handbook of Lifespan Development* (online). New York: Wiley.

Jensen, B. T. (2007). The relationship between Spanish use in the classroom and the mathematics achievement of Spanish-speaking kindergarteners. Journal of Latinos and Education 6(3), 267-280.

Jones, D. E., Greeberg, M., & Crowlety, M. (2015). Early social-emotional functioning and public health: The relationship between kindergarten social competence and future wellness. *American Journal of Public Health*.

Jordan, N. C., Hanich, L., & Kaplan, D. (2003). A longitudinal study of mathematical competencies in children with specific mathematics difficulties versus children with co-morbid mathematics and reading Difficulties. *Child Development*, *74(3*), 834-850.

Juel. C. (1988). Learning how to read and write; A longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology, 80*, 437-447.

Kallery, M. (2004). Early years teachers' late concerns and perceived needs in science: An exploratory study. Eurpean Journal of Teacher Education, 27(2), 147-165.

Kanno, Y., & Harklau, L. (2012). *Linguistic minority students go to college: Preparation, access, and persistence*. New York, NY: Routledge.

Kaplan, D. (2002). Methodological advances in the analysis of individual growth with relevance to education policy. *Peabody Journal of Education, 77*, 189-215.

Kimbrough Oller, D., & Eilers, R. E. (2002). *Language and literacy in bilingual children.* Clevedon, UK: Multilingual Matters.

Klein, S., Bugarin, R., Beltranena, R., & McArthur, E. (2004). Language minorities and their educational and labor market indicators: Recent trends. NCES 2004-009. Retrieved from National Center for Education Statistics: http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2004009

Kline, R. B. (1998). Principles and practice of structural equation modeling. New York, NY: Guilford Press.

Kline, R. B. (2011). Principles and practice of structural equation modeling (3rd ed.). New York, NY: Guilford Press.

Kormi-Nouri, R., Moniri, S., & Nilsson, L. (2003). Episodic and semantic memory in bilingual and monolingual children. *Scandinavian Journal of Psychology*, 44(1), 47-54.

Krashen, S., & McField, G. (2005). What works? Reviewing the latest evidence on bilingual education. *Language Learner, 1*(2), 7-10, 34.

Lin, Y.-C., Morgan, P.L., Farkas, G., Hillemeier, M., M., Cook, M., & Maczuga, S. (2013). Reading, mathematics, and behavioral difficulties interrelate: Evidence from a cross-lagged panel design and population-based sample of U.S. upper elementary students. *Behavioral Disorders*, *38*, 193-200.

Lindholm-Leary, K. J. (2005). The rich promise of two-way immersion. *Educational Leadership*, *62*(4), 56-59.

Lindholm-Leary, K. J. & Borsato, G. (2001). Impact of two way bilingual elementary programs on students' attitudes toward school and college. *Center for Research on Education, Diversity & Excellence*, *39*(1), 21-31.

Lindholm-Leary, K. & Genesee, F. (2010). Alternative educational programs for English language learners. In California Department of Education (Eds.), *Improving education for English language learners: Research-based approaches*, pp. 323-382. Sacramento: CDE Press.

Lindholm-Leary, K. & Genesee, F. (2014). Bilingual and biliteracy skills in young Spanish-speaking low-SES children: Impact of instructional language and primary language proficiency. *International Journal of Bilingual Education and Bilingualism*, *17*(2), 144-159.

Lindholm-Leary, K. J. & Howard, E. R. (2008). Language development and academic achievement in two-way bilingual elementary programs on students' attitudes toward school and college. In T. W. Fortune and D. J. Tedick (Eds.) (2008). Pathways to multilingualism: Evolving perspectives on immersion education (177-200). Clevedon, UK: Multilingual Matters.

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.

Lonigan, C. J., Allan, D. M., Goodrich, J. M., Farrington, A. L., & Phillips, B. M. (2015). Inhibitory control of Spanish-speaking language-minority preschool children measurement and association with language, literacy, and math skills. Journal of Learning Disabilities[Advance online publication].

Lopez, M., Barrueco, S., & Miles, K. (2006). *Latino infants and families: A national perspective of protective and risk factors for development*. Tempe, AZ: The National Task Force on Early Childhood Education for Hispanics.

Luk, G., DeSa, E. , & Bialystok, E. (2011). Is there a relation between onset age of bilingualism and enhancement of cognitive control? *Bilingualism*: Language and Cognition, *14*, 588-595.

Macias, R. F. (2014). Spanish as the second national language of the United States. *Review of Research in Education*, *38*(1), 33-57.

Macias, R. F. (2016). Language ideologies and rhetorical structures in bilingual education policy and research: Richard Ruiz's 1984 discursive turn. *Bilingual Research Journal*, 39(3-4), 173-179.

MacSwan, J., Thompson, M. S., deKlerk, G., McAlister, K. (2007). Theory, research, and theory-driven research: Explanations of academic achievement differences among English language learners. Paper presented at the University of California Linguistic Minority Research Institute, Arizona State University, Tempe, May 2-5, 2007.

Mallikarjun, A., Newman, R. S., & Novick, J. M. (2017). Exploiting the interconnected lexicon: Bootstrapping English languge learning in young Spanish speakers. *Transnational Issues in Psychological Science, 3*(1), 34-47.

Marcovitch, S., & Zelazo, P. D. (2009). A hierarchical competing systems model of the emergence and early development of executive function (Target article with commentaries). *Developmental Science, 12,* 1-18.

Marian, V., Shook, A., & Schroeder, S. R. (2013). Bilingual two-way immersion programs benefit academic achievement. *Bilingual Research Journal*, *36*(2), 167-186.

Martin, M., Mullis, I., & Foy, P. (2008). TIMSS 2007 international mathematics report. Findings from IEA’s Trends in International Mathematics and Science Study at the fourth and eighth grades, 319-323. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Marx, R. W., Bluemenfeld, P. C., Krajcik, J. S., Fishman, B., Soloway, E., Geier, R., & Revital, T. T. (2004). Inquiry-based science in the middle grades: Assessment of learning in urban systemic reform. *Journal of Research in Science Teaching*, *41*(10), 1063-1080.

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

MacCallum, R. C., and Hong, F. (1996). Power analysis and determination of sample size in covariance structure modeling. *Psychological Methods*, 1, 130-149.

McClelland, M. M., Cameron, C. E., Connor, C. M., Farris, C. L., Jewkes, A. M., & Morrison, F. J. (2007). Links between behavioral regulation and preschoolers' literacy, vocabulary, and math skills. *Developmental Psychology*, *43*, 947-959.

McClelland, M. M., Cameron, C. E., Wanless, S. B., & Murray, A. (2007). Executive function, behavioral self-regulation, and social-emotional competence. In O. N. Saracho, & B. Spodek (Eds.). *Contemporary perspectives on social learning in early childhood education* (pp. 83-107). Greenwich, CT: Information Age Publishing.

McClelland, M. M., Morrison, F. J. & Holmes, D. L. (2000). Children at risk for early academic problems: The role of learning-related social skills. *Early Childhood Research Quarterly*, *15*(3), 307-329.

McFarland, J., Hussar, B., Wang, X., Zhang, J., Wang, K., Rathbun, A., Barmer, A., Forrest Cataldi, E., and Bullock Mann, F. (2018). The Condition of Education 2018 (NCES 2018-144). U. S. Department of Education. Washington, DC: National Center for Education Statistics.
Retrieved [May 18, 2018] from https://nces.gov/pubsearch/pubsinfo.

Mendez, L. I., Crais, E. R., Castro, D. C., Kainz, K. (2015). A culturally and linguistic responsive vocabulary approach for young Latino dual language learners. *Journal of Speech and Language Hearing Research*, *58*(1), 93-106.

Merrill, C., & Daugherty, J (2010). STEM education and leadership: A mathematics and science partnership approach. *Journal of Technology Education*, *21*(2), 21-34.

Mertler, C., & Vannatta, R. (2013). *Advanced and Multivariate Statistical Methods, 5th ed*. Pyrczak Publishers.

Meuwissen, A. S., & Zelazo, P. D. (2014). Hot and cool executive function: Foundations for learning and healthy development. *Zero to Three, 35*(2), 18-23.

Miller, D. C., Sen, A., Malley, L. B., & Burns, S. D. (2009). Comparative indicators of Education in the United States and Other G-8 Countries: 2009 (NCES 2009-039). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U. S. Department of Education.

Minner, D. D., Levy, A. J., & Century, J. (2010). Inquiry-based science instruction-what is it and does it matter? Results from a research synthesis years 1984 to 2002. *Journal of Research in Science Teaching*, *41*(10), 1063-1080.

Moiduddin, E., Aikens, N., Tarullo, L., West, J., & Xue, Y. (2012). Child outcomes and classroom quality in FACES 2009. OPRE Report 2012-37aWashington, DC: Office of Planning, Research and Evaluation, Administration for Children & Families, U. S. Department of Health and Human Services [Retrieved from http://www.acf.hhs.gov/sites/default/files/opre/faces_2009.pdf]

Montanari, S. (2014). A case study of bi-literacy development among children enrolled in an Italian-English dual language program in Southern California. *International Journal of Bilingual Education and Bilingualism, 17*(5), 509-525.

Moorehead, T. & Grillo, K. (2013). Celebrating the reality of inclusive STEM education: Co-teaching in science and mathematics *Teaching Exceptional Children*, *45*(4), 50-57.

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., & Maczuga. (2012). Do poor readers feel angry, sad, and unpopular? *Scientific Studies of Reading, 16,* 360-381.

Morgan, P.L., Farkas, G., Tufis, P.A., & Sperling, R. A. (2008). Are reading and behavior problems risk factors for each other? *Journal of Learning Disabilities, 41,* 417-436.

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

Mulligan, G.M., Hastedt, S., & McCarrol, J.C. & National Center for Education Statistics (2012). *First time kindergartners in 2010-11: First findings from the kindergarten rounds of the early childhood longitudinal study, kindergarten class of 2010-11 (ECLS-K:2011).* (NCES 2012-049). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Mulligan, G. M., McCarroll, J. C., Flanagan, K. D., and Potter, D. (2014). *Findings from the first-grade rounds of early childhood longitudinal study, kindergarten class of 2010-11 (ECLS-K:2011)*. National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Mulligan, G. M., McCarroll, J. C., Flanagan, K. D., and Potter, D. (2015). *Findings from the second-grade rounds of early childhood longitudinal study, kindergarten class of 2010-11 (ECLS-K:2011)* (NCES 2015-077). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Mulligan, G. M., McCarroll, J. C., Flanagan, K. D., Potter, D., & National Center for Education

Statistics, & American Institutes for Research. (2016). *Findings from the third-grade round of the early childhood longitudinal study, kindergarten class of 2010-2011 (ECLS-K:2011)* (NCES 2016-094). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Muthén, B. O. (2002). Beyond *SEM*: General latent variable modeling. *Behaviormetrika*, 29, 1, 81-117.

Muthén, L. K., & Muthén, B. O. (1998-2011). Mplus User's Guide. Sixth Edition. Los Angeles, CA: Muthen & Muthen.

National Center for Education Statistics (2011). *The Nation's Report Card: Science 2009* (NCES 2011-451). Institute of Education Sciences, U.S. Department of Education, Washington, D. C.

National Center for Education Statistics. (2014). The condition on education 2014. Washington, D. C. Available at http://nces.ed.gov/programs/coe/

National Center for Education Statistics (2017). *Early childhood longitudinal program (ECLS) child assessments*. https://nces.ed.gov/ecls/assessments2011.asp

National Council for Teachers of Mathematics (1999). Principles and standards for school mathematics. Reston, VA: Author.

National Mathematics Advisory Panel (2008). *Foundations for success: The final report of the National Mathematics Advisory Panel.* Washington, DC: United State Department of Education.

National Research Council (2011). *Successful K-12 STEM education: Identifying effective approaches in science, technology, engineering, and mathematics*. Committee on Highly

Sucessful Science Programs for K-12 Science Education. Board on Science Education and Board on Testing and Assessment, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

National Science Foundation. (1994). Women, minorities, and persons with disabilities in science and engineering: 1994.

National Science Foundation, National Center for Science and Engineering Statistics (2011). *Women, minorities, and persons with disabilities in science and engineering: 2013.* Special Report NSF 15-311. Arlington, VA. Retrieved from http://www.nsf.gov/statistics/wmpd/archives/wmpd_2011.zip

National Science Foundation, National Center for Science and Engineering Statistics (2013). *Women, minorities, and persons with disabilities in science and engineering*: 2013. Special Report NSF 13-304. Arlington, VA. Retrieved from http://www.nsf.gov/statistics/wmpd

National Science Teachers Association (2014). Statement of early childhood science education. Retrieved from http://www.nsta.org/about/positions/earlychildhood.aspx.

Neathery, M. F. (1997). Elementary and secondary students' perceptions toward science and the correlation with the gender, ethnicity, ability, grade, and science achievement. *Electronic Journal of Science*, *2*(1), [online]. Available: http://unr.edu/homepage/jcannon/ejse/neathery.html.

New, R. S. (1999). Playing fair and square: Issues of equity in preschool mathematics, science, and technology. In: Dialogue on Early Childhood Science, Mathematics, and Technology Education: Fostering High Quality Programs. Washington DC: American Association for the Advancement of Science. Available at http://www.project2061.org/publications/earlychildhood/online/fostering/new.htm

Newfield, C. (2008). *Unmaking the public university: The forty-year assault on the middle class*. Cambridge, MA: Harvard University Press.

Organization for Economic Co-operation and Development. (2009). PISA *data analysis manual: SPSS, second edition*. Paris.

Organization for Economic Co-operation and Development. (2015). *PISA 2006 technical report*. Paris.

Osborne, J., Simon, S., & Collins, S. (2003). Attitudes towards science: A review of the literature and its implications. *International Journal of Science Education*, *25*(9), 1049-1079.

Parette, H., Quensenberry, A., & Blum, C. (2010). Missing the boat with technology usage in e early childhood settings: A 21st century view of developmentally appropriate practice. *Early Childhood Education Journal*, *37*, 335-343.

Park, M, H., Dimitrov, D. M., Patterson, L. G., & Park, D. Y. (2017). Early childhood teachers' beliefs about readiness for teaching science, technology, engineering, and mathematics. *Journal of Early Childhood Research*, *15*(3), 275-291.

Parsons, S., & Bynner, J. (1997). Numeracy and employment. *Education & Training, 39*, 43-51.

Passolunghi, M. C., & Siegal, L. S. (2001). Short-term memory, working memory, and inhibitory control in children with specific arithmetic learning disabilities. *Journal of Experimental Child Psychology*, *80*, 44-57.

Patrick, H., Mantzicopoulos, P., Samarapungavan, A., & French, B. F. (2008). Patterns of young children's motivation for science and teacher-child relationship. *The Journal of*

Experimental Education, *76*(2), 121-144.

Pham, A. V., & Hasson, R.M. (2014). Verbal and visuospatial working memory as predictors of children's reading ability. *Archives of Clinical Neuropsychology, 29*, 467-477.

Pickering, S. J., & Gathercole, S. E. (2004). Distinctive working memory profiles in children with special education needs. *Educational Psychology*, *24*, 393-408.

President's Council of Advisors on Science and Technology (2010). *Prepare and inspire: K-12 education in science, technology, engineering, and math (stem) for America's future.* Retrieved from The White House Issues Webpage: https://www.whitehouse.gov/sites/default/files/microsites/ostp/pcast-stemed-report.pdf

Price, L. R. (2017). *Psychometric methods: Theory into practice*. New York: The Guilford Press.

Provasnik, S., Malley, L., Stephens, M., Landeros, K., Perkins, R., and Tang, J. H. (2016). Highlights from TIMSS and TIMSS advanced 2015: Mathematics and science and science achievement of U.S. students in grades 4 and 8 and in advanced courses at the end of high school in an international context (NCES 2017-002). U.S. Department of Education, National Center for Education Statistics. Washington, D.C. Retrieved January 17, 2018 from http://nces.ed.gov/pubseratc

Przymus, S. D. (2016). Challenging the monolingual paradigm in secondary dual-language instruction: Reducing language-as-problem with the 2-1-L2 model. *Bilingual Research Journal*, *39*(3-4), 279-295.

Ramey-Gassert, L. (1997). Learning science beyond the classroom. *The Elementary School Journal* , 97(4), 433-450.

Ratey, J. (2002). *A user's guide to the brain: perception, attention, and the four theaters of the brain.* New York: Random House, Inc.

Reardon, S. (2003). Sources of educational inequality: The growth of racial/ethnic and socioeconomic test score gaps in kindergarten and first grade. Population Research Institute, Pennsylvania State University, University Park. PA.

Reardon, S., & Galindo, C. (2006). K-3 academic achievement patterns and trajectories of Hispanics and other racial/ethnic groups. Paper presented at the Annual American Education Research Association Conference, San Francisco, CA.

Rivera-Batiz, F. (1992). Quantitative literacy and the likelihood of employment among young adults in the United States. *The Journal of Human Resources*, 27, 313-328.

Roberts, G. & Bryant, D. (2011). Early mathematics achievement trajectories: English language learner and native English-speaker estimates, using the early childhood longitudinal survey. *Developmental Psychology, 47(4*), 916-930.

Rodriguez, G. M., & Cruz, L. (2009). The transition to college of English Learner and undocumented immigrant students: Resource and policy implications. *Teachers College Record*, *111*(10), 2385-2418.

Rolstad, K., Mahoney, K., & Glass, G. V. (2005). The big picture: A meta-analysis of program effectiveness research on English language learners. *Educational Policy*, 19(4), 572-594.

Rossell,, C. H. (1990). The effectiveness of educational alternatives for limited-English proficient children. In G. Imhoff (Ed.), *Learning in two languages* (pp. 71-121). New Brunswick, NJ: Transaction Publishers.

Rossell, C. H., & Baker, K. (1996). The educational effectiveness of bilingual education.

Research in the Teaching of English, *30*(1), 7-74.

Rothlisberger, M., Neuenschwander, R., Cimeli, P., & Roebers, C. M. (2013). Executive functions in 5- to 8-year- olds: Developmental changes and relationship to academic achievement. *Journal of Educational & Developmental Psychology*, *3*, 153-167.

Rubinstein-Avila, E. (2002). Problematizing the "dual" in a dual-immersion program: A portrait. *Linguistics and Education*, 13(1), 65-87.

Rumbaut, R. G. (2005). Turning points in the transition to adulthood: Determinants of educational attainment, incarceration, and early childbearing among children of immigrants. *Ethnic and Racial Studies*, *28*(6), 1041-1086.

Saunders, W. M., & Marcelletii, D. J. (2013). The gap that can't go away: The catch-22 of reclassification in monitoring the progress of English learners. Educational Evaluation and Policy Analysis, 35(2): 139-156. Retrieved September 28, 2017, from http://journals.sagepub.com/doi.full/10.3102/0162373712461849.

Senn, T., Espy, K., & Kauffmann, P. (2004). Using path analysis to understand executive function and organization in preschool children. *Developmental Neuropsychology*, *26*, 445-464.

Shin, T., Davison, M., Long, J., Chan, C., & Heistad, D. (2013). Exploring gains in reading and mathematics achievement among regular and exceptional students using growth curve modeling. *Learning and Individual Differences*, *23*, 92-100.

Samuels, W. E., Tournaki, N., Blackmon, S. & Zilinski, C. (2016). Executive functioning predicts academic achievement in middle school: A four-year longitudinal study. *Journal of Educational Rresearch*, 109, 5, 478-490.

Shneyderman, A., & Abella, R. (2009). The effects of the extended foreign language problems on Spanish-language proficiency and academic achievement in English. Bilingual Research Journal, 32(3), 241-259.

Slavin, R. E., & Cheung, A. (2005). A synthesis of research on language of reading instruction. *Review of Educational Research*, *75*(2), 247-284.

Slavin, R. E., Madden, N. A., Calderon, M., Chamberlain, A., & Hennessy, M. (2011). Reading and language outcomes of a multiyear randomized evaluation of transitional bilingual education. *Education Evaluation and Policy Analysis*, 33(1), 47-58.

Snyder, T. D., de Brey, C., & Dillow, S. A. (2016). *Digest of Education Statistics* (NCES 2016-014). National Center for Education Statistics, Institute of Education Sciences, U. S. Department of Education. Washington, D. C.

Soderman, A. (2010). Language immersion program for young children? Yes…but proceed with caution: Young children can benefit from dual-language immersion programs that are developmentally appropriate. *Phi Delta Kappan*, 91(8), 51-59.

Steele, J. s., Slater, R. O., Zamarro, G., Miller, T., Li, J., Burkhauser, S., & Bacon, M. (2017). Effects of dual-language immersion program on student achievement: Evidence from l lottery data. *American Educational Research Journal*, *54*(1S), 282S-306S.

Stepler, R., & Lopez, M. H. (2016). U.S. Latino population growth and dispersion has slowed since onset of the great recession. Pew Research Center, Washington, D.C. Retried from: http://www.pewhispanic.org/2016/09/08/latino-population-growth-and-dispersion-has-slowed-since-the-onset-of-the-great-recession/

Stewart, J. H. (2005). Foreign language study in elementary schools: Benefits and implications for achievement in reading and math. *Early Childhood Education Journal, 33*(1), 11-16.

Strutchens, M. E. & Silver, E. A. (2000). NAEP findings regarding race/ethnicity: The students, their performance, and their classrooms. In E. A. Silver & P. A. Kenney (Eds.), *Results from the seventh mathematics assessment of the National Assessment of Educational Progress* (pp.45-72). Reston, VA: National Council of Teachers of Mathematics.

Successful STEM Education (2013). Nurturing STEM skills in young learners, PreK-3. Retrieved from http://www.successfulstemeducation.org/resources/briefs.

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

Swanson, H. L., & Saez, L. (2003). Memory difficulties in children and adults with learning disabilities. In H. L. Swanson, K. R. Harris & S. Graham (Eds.). Handbook of learning disabilities (pp.182-198).

Swanson, H. L., Zheng, X., & Jerman, O. (2009). Working memory, short-term memory, and reading disabilities. *Journal of Learning Disabilities, 44*, 260-287.

Tabachnick, B. G. & Fidell, L. S. (2012). Using Multivariate Statistics, 6th ed. Pearson: Allyn & Bacon Publishers.

Tarr, J. E., Reys, R. E., Reys, B. J., Chavez, O., Shih, J., & Osterlind, S. J. (2008). The impact of middle-grades mathematics curricula and the classroom learning environment on student achievement. *Journal for Research in Mathematics Education*, *39*, 247-280.

The Nations Report Card: Mathematics 2009 (2010). National Assessment of Educational Progress at Grades 4 and 8 (NCES 2010-451). Washington, DC: National Center for Education Statistics, U. S. Department of Education.

Thomas, W., & Collier, V. P. (2002). A national study of school effectiveness for language minority students' long term academic achievement (Final report: Project 1.1). Santa Cruz: University of California, Center for Research on Education, Diversity, and Excellence.

Thomas, W. P., & Collier, V. P. (2004). The astounding effectiveness of dual language education for all. *NABE Journal of Research and Practice*, *2*(1), 1-20.

Tochon, F. V. (2009). The key to global understanding: World languages education-why schools need to adapt. *Review of Educational Research*, *79*(2), 650-681.

Toll, S.W. M., van der Ven, S., Kroesbergen, E., & van Luit, J. (2011). Executive functions as predictors of math learning disabilities. *Journal of Learning Disabilities*, 44, 521-532.

Tominey, S. L., and McClelland, M. M. (2011). Red light, purple light: Findings from a randomized trial using circle time games to improve

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.

Tourangeau, K., Nord, C., Lê, T., Wallner-Allen, K., Hagedorn, M. C., Leggitt, J., & Westat, I. (2015). *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 2015-078). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Tourangeau, K., Nord, C., Lê, T., Wallner-Allen, K., Hagedorn, M. C., Leggitt, J., & Westat, I. (2015). *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 2015-078). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Tourangeau, K., Nord, C., Lê, T., Wallner-Allen, K., Hagedorn, M. C., Leggitt, J., & Najarian, M. (2014). *Early Childhood Longitudinal Study, Kindergarten Class of 2010-11 (ECLS-K:2011). User's Manual for the ECLS-K:2011 Kindergarten-First Grade Data File and Electronic Codebook, Public Version*. NCES 2015-069). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Tourangeau, K., Nord, C., Lê, T., Sorongon, A. G., Hagedorn, M. C., Daly, P., & Najarian, M. (2013). *Early Childhood Longitudinal Study, Kindergarten Class of 2010-11 (ECLS-K:2011). User's Manual for the ECLS-K:2011 Kindergarten Data File and Electronic Codebook, Public Version*. NCES 2013-061). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Tran, N. A., Martinez-Cruz, A., Behseta, S., Ellis, M., Contreras, J. (2015). The effects of Spanish English dual language immersion on student achievement in science and mathematics. *eJournal of Education Policy*, 57-77.

Trochim, W. M. K. (2006). Positivism and post-positivism. Retrieved from Research Methods Knowledge Base website, http://www.socialresearchmethods.net/kb/positivsm.php

U. S. Census Bureau. (2010). Overview of race and Hispanic origin. Retrieved from http://www.census.gov/prod/cen2010/briefs

United States White House Issues Educate to Innovate. (2015). Retrieved from The White House Issues Website: https://www.whitehouse.gov/issues/education/k-12/educate-innovate.

Van der Ven, S. H. G., Kroesbergen, E. H. Boom, J., & Leseman, P.P.M. (2012). The development of executive functions and early mathematics: A dynamic relationship. *British Journal of Educational Psychology*. *82*, 100-119.

Van der Ven, S. H., Boom, J., Kroesbergen, E. H., P.P. Leseman (2012). Microgenetic patterns of children's multiplication learning: Confirming the overlapping waves model by latent growth modeling. *Journal of Experimental Child Psychology*, 113 (2012), pp. 1-19

Vega, Clemente III. (2008). Cognitive flexibility advantages in bilingual children. Available from ProQuest Dissertations & Theses Global. (304820398). Retrieved from http://libproxy.txstate.edu/login?url=https://search-proquest-com.libproxy.txstate.edu/docview/304820398?accountid=5683

Viterbori, P., Usai, M. C., Traverso, L., & De Franchis, V. (2015). How preschool executive functioning predicts several aspects of math achievement in grades 1 and 3: A longitudinal study. *Journal of Experimental Child Psychology*, *140*, 38-55.

Waber, D. P., Gerber, E. B., Turcios, V. Y., Wagner, E. R. & Forbes, P. W. (2006). Executive functions and performance on high-stakes testing in children form urban schools. *Developmental Neuropsychology*, *29*(3), 459-477.

Watters, J. J., Diezmann, C. M., Grieshaber, S. J., & Davis, J. M. (200). Enhancing science education for young children: A contemporary initiative. *Australian Journal of Early*

Childhood, *26*(2), 1-7.

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*, 43-53.

West, S. G., Finch, J. F., & Curran, P. J. (1995). Structural equation models with nonnormal variables: Problems and remedies. In R. H. Hoyle (Ed.), Structural equation modeling: Concepts, issues, and applications (pp.56-75). Thousand Oaks, CA: Sage Publications.

Wiebe, S. A., Sheffield, T., Neslon, J. M., Clark, C. A. C., Chevalier, N., & Espy, K. (2011). The structure of executive function in 3-year-olds. Journal of Experimental Child Psyhology, 108, 43-452.

Wiley, T. G. (2014). Diversity, super-diversity, and monolingual language ideology in the United States: Tolerance or intolerance? *Review of Research in Education*, *38*(1), 1-32.

Wiley, T. G., & Garcia, O. (2016). Language policy and planning in language education legacies, Consequences and possibilities. *The Modern Language Journal*, *100*(S1), 48-63.

Wiley, T. G., & Wright, W. E. (2004). Against the undertow: Language-minority education policy and politics in the "age of accountability." *Educational Policy*, 18(1), 142-168.

Zelazo, P. D. (2006). The dimensional change card sort (DCCS): A method of assessing executive function in children. *Nature Protocols, 1*, 297-301.

Zelazo, P.D., Anderson, J. E., Richler, J., Wallner-Allen, K., Beaumont, J. L., and Weintraub, S. (2013). NIH Toolbox Cognition Battery (CB): Measuring executive function and attention. In P. D. Zelazo and P. J. Bauer (Eds.), *National Institutes of Health Toolbox*

Cognition Battery (NIH Toolbox CB): Validation for children between 3 and 15 years. *Monographs for the Society for Research in Child Development.*

Zelazo, P. D., Blair, C. B., & Willoughby, M. T. (2017). Executive function: Implications for Education NCER 2017-2000. Washington, DC: National Center for Education Research, Institute US Department of Education, Institute of Education Sciences, US. Department of Education. Available at http://ies.ed.gov

Zelazo, P. D., Carter, A., Reznick, J. S., & Frye, D. (1997). Early development of executive function: A problem-solving framework. *Review of General Psychology*, 1(2), 198-226.

Zhang, Z., McArdle, J. J., Nesselroade, J. R. (2012). Growth rate models: Emphasizing growth rate analysis through growth curve modeling. *Journal of Applied Statistics, 39(6)*, 1241-1262.

Zuniga, C. E. (2016). Between language as problem and resource: Examining teachers' language orientations in dual-language programs. Bilingual Research Journal, 39(3/4), 339-353.