1 + 1 IS NOT ALWAYS 2: VARIATION IN THE RELATIONS BETWEEN MATHEMATICS SELF-EFFICACY DEVELOPMENT AND LONGITUDINAL MATHEMATICS ACHIEVEMENT GROWTH

by

CAROLINE JANE SHANLEY

A DISSERTATION

Presented to the Department of Educational Methodology, Policy, and Leadership
and the Graduate School of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

December 2014

DISSERTATION APPROVAL PAGE

Student: Caroline Jane Shanley

Title: 1 + 1 Is Not Always 2: Variation in the Relations Between Mathematics Self-Efficacy Development and Longitudinal Mathematics Achievement Growth

This dissertation has been accepted and approved in partial fulfillment of the requirements for the Doctor of Philosophy degree in the Department of Educational Methodology, Policy, and Leadership by:

Gina Biancarosa	Chairperson
Ben Clarke	Core Member
Mark Van Ryzin	Core Member
Joanna Goode	Institutional Representative

and

J. Andrew Berglund	Dean of the Graduate School

Original approval signatures are on file with the University of Oregon Graduate School.

Degree awarded December 2014

DISSERTATION ABSTRACT

Caroline Jane Shanley

Doctor of Philosophy

Department of Educational Methodology, Policy, and Leadership

December 2014

Title: 1 + 1 Is Not Always 2: Variation in the Relations Between Mathematics Self-Efficacy Development and Longitudinal Mathematics Achievement Growth

Creating an educational program that results in positive post-secondary and science, technology, engineering, and mathematics (STEM)-oriented outcomes for all students is a national goal and federal policy directive. Recent research has shown that in addition to measures of academic proficiency, intra- and interpersonal skills are important factors in college and career readiness. Likewise, mathematics proficiency is an important skill for successful STEM outcomes and post-secondary success, but these achievements and outcomes frequently vary based on demographic characteristics. This study utilized data from the Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 to examine the relationships between mathematics achievement growth in Grades K–1 and Grades 3–8, mathematics self-efficacy development in Grades 3–8, and demographic factors including sex, socioeconomic status (SES), and race/ethnicity. Various models of mathematics achievement growth were tested, and the relationships between both early and middle grades mathematics achievement growth and self-efficacy development were also explored. Sex, SES, and race/ethnicity differences in both mathematics achievement growth and self-efficacy development were discovered, and findings were consistent with familiar achievement gaps favoring white and Asian males from above median SES households. In particular, SES was found to be a ubiquitous factor in both mathematics achievement and self-efficacy development, and sex moderated some of the relationships between mathematics achievement and self-efficacy. Implications for future research, instructional design, and intervention development are discussed.

CURRICULUM VITAE

NAME OF AUTHOR: Caroline Jane Shanley

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene, OR
San Francisco State University, San Francisco, CA
Whitman College, Walla Walla, WA

DEGREES AWARDED:

Doctor of Philosophy, Educational Leadership, 2014, University of Oregon
Master of Arts, Special Education, 2011, San Francisco State University
Bachelor of Arts, History, 2000, Whitman College

AREAS OF SPECIAL INTEREST:

Mathematics Education

Research Design & Quantitative Methodology

PROFESSIONAL EXPERIENCE:

Doctoral Fellow; Lead Author & UTFG Lead; KinderTEK Intervention Author; Cognitive Study Co-Coordinator, University of Oregon, Center for Teaching and Learning (CTL), 2011-present

Early Elementary Mathematics Content Expert; Item Reviewer, Southern Methodist University, Research in Mathematics Education, 2013-present

Research & Analytics: Reliability Analyst, Program Administrator Intern University of Oregon, Educational Policy Improvement Center (EPIC), 2012 & 2014

Resource Specialist, Mild/Moderate Disabilities, K-12, Oakland Unified School District, Oakland, CA, 2000-2011

GRANTS, AWARDS, AND HONORS:

Educational Methodology, Policy, and Leadership Department Travel Grant Award, 2013 & 2014

Graduate Teaching Fellow, University of Oregon, Center on Teaching and Learning, 2013-2014

Statistics Institute on Mathematics Equity Participant Award, American Education Research Association, 2012

Doctoral Research Fellowship, Center on Teaching and Learning, 2011-2013

PUBLICATIONS:

Clarke, B., Doabler, C. D., Nelson-Walker, N., & Shanley, L. (in-press). Effective early numeracy and whole number concepts instruction for all learners: Translating research to practice. *Intervention in School & Clinic.*

Clarke, B., Nelson, N., & Shanley, L. (in-press). Mathematics fluency: more than the weekly timed test. In K.D. Cummings & Y. Petscher (Eds.). *Fluency metrics in education: Implications for test developers, researchers, and practitioners.* New York: Springer.

Biancarosa, G., & Shanley, L. (in-press). What is fluency? In K.D. Cummings & Y. Petscher (Eds.). *Fluency metrics in education: Implications for test developers, researchers, and practitioners.* New York: Springer.

Carlson, S. E., Biancarosa, G., Basaraba, D. L., & Shanley, L. (2014). Supporting readers in science. *ASCD Express, 9*(13).

Basaraba, D. L., Biancarosa, G., Carlson, S. E., & Shanley, L. (2014). Teaching students to read like historians. *ASCD Express, 9*(13).

ACKNOWLEDGMENTS

Many thanks to Dr. Gina Biancarosa for her assistance in preparation of this manuscript and for her invaluable mentorship and support. In addition, I would like to thank Dr. Ben Clarke for his support of this project and Drs. Mark Van Ryzin and Joanna Goode for their thoughtful input.

For my dad, who taught me that attitude is everything; for Ryan, who always makes me think; and for my beautiful biscuit, whose brilliance and potential is unmatched.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Table Page

CHAPTER I

INTRODUCTION

Demonstrating proficiency in mathematics is an important prerequisite for many academic disciplines, and mathematics achievement is a critical predictor of future academic attainment (Gamoran & Hannigan, 2000; Meltzer, 2002). Consequently, standards developers and policy makers have advocated for the implementation of early, targeted mathematics instruction and supports (Common Core State Standards Initiative, 2010). In the federal blueprint for the reauthorization of the Elementary and Secondary Education Act (ESEA; U.S. Department of Education, 2010), the U.S. Department of Education called for the institution of rigorous academic standards that emphasized Science, Technology, Engineering, and Mathematics (STEM) education. Additionally, the Common Core State Standards Initiative (2010) specifies a set of learning objectives to ensure that all students develop comprehensive mathematics understandings in a range of content areas (e.g., Operations & Algebraic Thinking, Number & Operations in Base Ten, Geometry, etc…) beginning as early as kindergarten. Recent standards movements have worked to develop instructional foci organized around a coherent progression of mathematics standards to prepare students for a range of future academic and career endeavors.

Also essential factors in supporting mathematics achievement, intra and interpersonal skills (i.e. self-efficacy, academic self-confidence, self-appraisal, resourcefulness, perseverance; Lee & Stankov, 2013) work in concert with academic skills to facilitate positive longitudinal outcomes (Dweck, Walton, & Cohen, 2011). In fact, some researchers contend that much of the long-term academic success attributed to early intervention programs can actually be credited to program features that increase intra and interpersonal skill development (Heckman, 2006; Heckman & Rubinstein, 2001). Identifying the relationships between trajectories of academic achievement growth and the development of critical intra and interpersonal skills, and the extent to which those relationships hold across demographic categories may have ramifications for educational planning and instruction as early as kindergarten; however, these relationships are currently largely unknown.

This study represents an attempt to inform academic planning efforts and intervention development through an analysis of patterns of mathematics achievement growth and mathematics self-efficacy development across the elementary and middle school grades in a diverse sample of learners in the Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K) dataset. The associations between features of Grades K–8 mathematics growth trajectories and mathematics self-efficacy growth parameters in Grades 3–8 are evaluated, and the moderating effect of demographic factors are tested. Identifying patterns of mathematics growth that are associated with efficacy in mathematics has implications for curriculum development and the implementation of instructional programs. Moreover, understanding how these relations may differ for different groups can inform academic planning so that resources are applied most effectively for all students to ensure that all students receive effective foundational mathematics instruction and develop intra and interpersonal skills that support future achievement.

CHAPTER II

LITERATURE REVIEW

Relationships between many of the factors considered here have robust research histories and are fairly widely known. For example, (a) early mathematics achievement predicts later mathematics achievement (Claessens & Engel, 2013; Duncan, et al., 2007; Morgan, Farkas, & Wu, 2009), (b) academic achievement and self-efficacy have a complementary relationship where high academic achievement is associated with high academic self-efficacy and academic difficulties are related to low reports of self-efficacy (Caprara, Vecchione, Alessandri, Gerbino, & Barbaranelli, 2011; Diseth, 2011), and (c) demographic characteristics are associated with educational outcomes and define a number of persistent achievement gaps (Burchinal et al., 2011; Nowell & Hedges, 1998; Reardon, 2011). However, the extent to which mathematics achievement growth is related to mathematics self-efficacy development and the ways in which the relationships between mathematics achievement growth and the development of mathematics self-efficacy differ for students based on sex, socioeconomic status (SES), and race/ethnicity remain relatively unexplored. The following sections provide an overview of what is known about each factor considered here and identify key questions that remain unanswered.

Mathematics Achievement

Although a variety of academic skills are required for postsecondary success, mathematics achievement is a particularly influential factor in future outcomes (Pellegrino & Hilton, 2012). For example, mathematics attainment in secondary school largely dictates college course taking and degree completion (Lee, 2012) and secondary mathematics achievement is indicative of career earnings (Achieve, 2008). Whereas early mathematics performance is known to predict later mathematics performance, a detailed exploration of the relationship between Grades K–1 and Grades 3–8 mathematics achievement growth has utility for identifying patterns of mathematical learning within and between each time frame. These understandings can inform mathematics instruction and the development of intervention programs across the elementary and middle school years.

Early grades. Research suggests that primary grade mathematics achievement may be particularly important for long-term success because kindergarten academic skills, measured upon school entry are highly predictive of academic performance at the end of first grade and well beyond. Numerous longitudinal studies of mathematics development have shown that deficits observed as early as kindergarten are difficult to overcome and often compound over time (Duncan, et al., 2007; Morgan, Farkas, & Wu, 2009), and mathematics learning in kindergarten and first grade is associated with later mathematics achievement **(**Claessens, Duncan, & Engel, 2009; Duncan & Murnane, 2011**).** In fact, one longitudinal study found that students who were in the 10th percentile for mathematics achievement at both entrance and exit from kindergarten had only a 30% chance of performing above the 10th percentile five years later (Morgan, Farkas, & Wu, 2009). Thus, it is hypothesized that students' later patterns of growth in mathematics achievement (i.e., Grades 3–8) are related to not only their level of knowledge or skill upon school entry (i.e., initial status in kindergarten), but also by how much they learned in kindergarten and first grade (i.e., rate of mathematics achievement growth in Grades K–1). Furthermore, if early mathematics learning is influential in later mathematics learning, Grades K–1 mathematics achievement growth may also be predictive of later mathematics self-efficacy development. Therefore, an additional hypothesis for this study is that Grades K–1 mathematics achievement growth has a unique, positive relationship with Grades 3–8 self-efficacy development even when controlling for Grades 3–8 mathematics achievement.

Middle grades. Along with the strong links between Grades K–1 mathematics achievement and later mathematics achievement, research has found that being prepared in the middle grades to pursue a rigorous high school mathematics curriculum is a robust predictor of later outcomes (Adelman, 2006). For example, eighth grade mathematics achievement scores are strongly associated with enrollment in advanced mathematics courses in high school and beyond (Gamoran & Hannigan, 2000) and the skills that lead to eighth grade mathematics achievement can be traced back to content knowledge taught in the late elementary and middle grades. In fact, a recent longitudinal study found that knowledge of division and of fractions was most predictive of algebra readiness and secondary mathematics achievement, even when controlling for general cognitive ability and demographic factors (Siegler et al., 2012). Because mathematics achievement in the middle grades is uniquely predictive of factors that support long-term academic achievement, it is hypothesized that Grades 3–8 mathematics growth trajectories will be positively associated with other longitudinal indicators (i.e., intra and interpersonal skills such as mathematics self-efficacy).

Self-Efficacy

Mastery of academic content, as measured by grades and test scores, is one indicator of future outcomes, but intra or interpersonal skills are at least equally important in academic retention, completion, and advancement (Camara, 2005; Farrington et al., 2012; Pellegrino & Hilton, 2012). Grit, determination, persistence, work habits, and self-efficacy are all individual characteristics that correlate with positive academic outcomes, and improving these intra and interpersonal skills can have a direct effect on student grades and other measures of academic performance (Dweck, Walton, & Cohen, 2011). For example, a recent meta-analysis of 213 school-based social and emotional intervention programs found that participating students demonstrated moderate gains (Cohen's $d = .27$) in academic performance in addition to social and emotional skill and behavior improvements (Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011).

Within the spectrum of influential intra and interpersonal skills, self-efficacy alone has been found to be an important indicator of academic achievement and retention (Bandura, 1986; Di Giunta et al., 2013; Mattern & Shaw, 2010; Pajares & Schunk, 2001). Self-efficacy refers to beliefs about one's capabilities to learn or perform behaviors at designated levels and is a flexible trait that develops in response to experience and achievement (Bandura, 1986). Academic self-efficacy and academic achievement have a reciprocal nature such that students who experience success in school are generally more likely to believe in their ability to perform academic tasks (Diseth, 2011) and students who report high levels of self-efficacy tend to be successful in school (Caprara et al., 2011). Additionally, self-efficacy is positively associated with both perseverance and resiliency (Pajares, 1996). In mathematics specifically, self-efficacy has been correlated both with other intra and interpersonal skills and with mathematics achievement (Shams, Mooghali, & Soleimanpour, 2011), and mathematics self-efficacy is an effective predictor of mathematics performance (Pajares & Miller, 1994) and the likelihood of pursuing STEM degrees (Larson et al., 2014). Furthermore, research on mathematics achievement gaps suggests that achievement gaps decrease when controlling for mathematics self-efficacy (Kitsantas, Cheema, & Ware 2011).

The emergence of mathematics achievement as a particularly important foundational skill and its close relationship with self-efficacy has prompted the need to evaluate the relationship between achievement growth trajectories and the development of or change in academic self-efficacy. Research has shown that self-efficacy beliefs often decrease through school, especially in the middle grades (Midgley, Feldlaufer, & Eccles, 1989), perhaps in response to competition, experiences of difficulty and/or failure, and the sometimes inflexible pace and environment of formal schooling (Pintrich & Schunk, 1996; Bandura, 1997). Of course, not all students experience this dip in academic self-efficacy and because there is evidence of a positive relationship between mathematics self-efficacy and mathematics achievement, a better understanding of whether or not there are critical developmental periods for both academics and self-efficacy (i.e., Grades K–1) and the relationships between specific academic subject growth (i.e., mathematics) and the development of self-efficacy is critical. For example, it could be that Grades K–1 mathematics performance uniquely predicts mathematics self-efficacy development in Grades 3–8 and early mathematics achievement is highly influential for later intra and interpersonal skill development. Alternatively, the relationship between Grades K–1 mathematics growth and later self-efficacy development could be mediated by more proximal mathematics achievement growth in Grades 3–8 and the concurrent development of academic and intra and interpersonal skills should be targeted and fostered in the middle grades.

Group Differences

In addition to flexible intra and interpersonal skills such as self-efficacy, traditionally fixed demographic factors (i.e., sex, SES, and race/ethnicity) are also associated with academic achievement growth and contribute greatly to the prediction of longitudinal outcomes. Sex, SES, and race/ethnicity differences in academic performance emerge as early as kindergarten and persist throughout elementary school and beyond (Duncan & Magnuson, 2005; Hair, Halle, Terry-Humen, Lavelle, & Calkins, 2006). Whereas these differences are typically framed as long-term achievement gaps, it is worthwhile to note that there is evidence that differences in academic performance between kindergarten entry and the end of first grade shrink for some students (Reardon & Galindo, 2009). Such findings suggest that the relationships between mathematics achievement, mathematics self-efficacy, and demographic characteristics are likely quite variable. Thus, because previous research suggests that factors that predict postsecondary success differ based on student demographic characteristics (Linver & Davis-Kean, 2005; Tracey & Robbins, 2004**)** and the presence distinct patterns of achievement based on sex, SES, and race/ethnicity has been documented (Long, Iatarola, & Conger, 2009; Tracey & Robbins, 2004)**,** a deeper investigation of whether mathematics achievement growth trajectories and mathematics self-efficacy development differ based on sex, SES, race/ethnicity and whether the relationship between mathematics achievement and mathematics self-efficacy development differed based on these same demographic characteristics is warranted.

Sex. Sex differences with regard to mathematics achievement growth and the development of mathematics self-efficacy may be in some part attributed to social cognitive theory where students view STEM subjects as male domains (Eisenberg, Martin, & Fabes, 1996; Nosek & Smyth, 2011). This sex achievement gap is evidenced by research findings suggesting that postsecondary outcomes and the likelihood of pursuing STEM careers vary based on sex (Duncan & Murnane, 2011). For example, women comprise more than 50% of the American college student population (Jacob, 2002) but less than 25% of all STEM related employment positions (Beede et al., 2011), and women are half as likely as men to choose a STEM related major in college (Morgan, Gelbgiser, & Weeden, 2013).

The sex gap appears to apply to intra and interpersonal skills, as well. Recent research conducted with elementary students suggests that the relationships between mathematics achievement and self-efficacy differ based on sex such that boys consistently report higher levels of mathematics self-efficacy regardless of academic performance (Joët, Usher, & Bressoux, 2011). However, there are a number of methodological issues and potentially confounding factors to consider. For example, boys tend to rate themselves higher than girls when given an absolute scale, but girls often rate themselves higher when asked to make comparative judgments of their abilities (Pajares, Miller, & Johnson, 1999). Additionally, girls have been found to provide more modest self-evaluations than boys (Wigfield, Eccles, & Pintrich, 1996) and these differences are also related to age and developmental levels (Eccles & Midgley, 1989).

Socioeconomic status (SES). The relationship between SES and academic achievement is well documented. A recent meta-analysis (Sirin, 2005) found that (a) the relationship between SES and academic achievement is moderate to strong and is slightly moderated by various contextual factors, and (b) there appeared to be a slight decrease in the correlation between SES and achievement compared to a similar meta-analysis conducted 20 years prior (see White, 1982). Specific study findings suggest that SES is a reliable predictor of early mathematics achievement as students from low income households tend to enter kindergarten with less developed mathematics skills and perform lower than their more affluent peers on standardized mathematics assessments (Crane, 1996; Lee & Burkam, 2002). Students from low-SES backgrounds also tend to have less access to college preparatory mathematics and science courses, are less likely to enroll in college or secure employment when they graduate from high school, and receive less support for postsecondary planning (Lippman, Burns, & McArthur, 1996; Wimberly & Noeth, 2005).

Consistent with the positive relationship between SES and academic achievement, SES is also closely related to self-efficacy. Students from low SES backgrounds tend to report lower levels of self-efficacy (Schunk & Meece, 2006). In turn, postsecondary research has shown that individuals who report higher levels of self-efficacy are more likely to attain higher SES levels in the form of higher job satisfaction, larger salaries, and more satisfactory occupational status (Judge & Hurst, 2008). Because SES is an indicator of access and opportunity, which often leads to experience, self-efficacy naturally follows based on its relationship with experience and success.

Race and ethnicity. Race and ethnicity-based achievement gaps are well documented and persist in spite of targeted intervention efforts. In fact, a 30-year examination of National Assessment of Educational Progress (NAEP) data showed that although racial and ethnic achievement gaps have shrunk slightly, they remain substantively significant and associated with inequitable educational opportunities

(Berends, Lucas, Sullivan, & Briggs, 2005; Kao & Thompson, 2003; Lee, 2004). For example, comparative analyses of Black-White mathematics achievement find that significant gaps remain even when controlling for sex and SES (Lubienski, 2002; Vanneman, Hamilton, Anderson, & Rahman, 2009). The long-term effects of discriminatory laws and policies remain present in both academic settings and post-secondary outcomes (Howard, 2010; Jencks & Phillips, 2011; Whaley & Noel, 2011).

Race and ethnicity differences have also been observed in reports of academic self-efficacy. A study conducted with a diverse sample of young adolescents found that reports of academic self-efficacy, in addition to the perceived ability to find a meaningful career, and possessing effective means for longitudinal goal attainment all differed based on race and ethnicity (Smith, Walker, Fields, Brookins, & Seay, 1999). Additionally, research suggests that the impact of intra and interpersonal skills on post-secondary pursuits may also differ based on race and ethnicity. Fortin (2008) found that intra and interpersonal skills were influential predictors of the Black-White career outcomes gap. However, a review of African American student achievement motivation demonstrated that African American students do not have lower perceptions of self-efficacy as compared to White students when controlling for SES (Graham, 1994) suggesting that race and ethnicity should be considered in concert with other relevant demographic factors.

Summary. Because there are clear patterns of achievement and post-secondary outcomes based on demographic characteristics, attending to how particular features of mathematics achievement growth measured at various points in time are related to the development of self-efficacy in mathematics and how the complex relations between mathematics achievement growth and mathematics self-efficacy development are predicted and moderated by demographic characteristics, is a vital activity for developing academic programs that prepare all students for longitudinal achievement. Previous studies found that sex, SES, eighth grade academic performance, and interactions between each of the factors are uniquely indicative of postsecondary educational choices (Trusty, Robinson, Plata, & Ng, 2000) and racial/ethnic and sex factors are interrelated (Perez-Felkner, McDonald, Schneider, & Grogan, 2012). Thus, it is hypothesized here that sex, SES, and race/ethnicity will be associated with both mathematics achievement growth and the development of mathematics self-efficacy with girls, students from low SES backgrounds, and students from traditionally underserved racial and ethnic backgrounds demonstrating lower levels of mathematics achievement growth and self-efficacy development. It is also hypothesized that the relationships between mathematics achievement growth and mathematics self-efficacy development will be moderated by demographic characteristics such that mathematics achievement growth may be less predictive of mathematics self-efficacy for female students, but more predictive of the development of mathematics self-efficacy for students who are not White and who come from low-SES backgrounds. Additionally, it is hypothesized that relationships demographic factors and mathematics achievement growth or self-efficacy development may attenuate in the presence of others. Specifically, because previous research suggests that differential academic outcomes and intra and interpersonal skill reports based on race were reduced when controlling for SES (Graham, 1994), the effects of both sex and race/ethnicity on self-efficacy and achievement may be less when controlling for SES. Considering these demographic factors jointly and modeling their relations with mathematics achievement growth and self-efficacy development in combination with one another is an important consideration to generate a more complete picture of the factors at play.

Current Study

This study had five main aims. The first aim was to model and evaluate Grades K-8 mathematics achievement growth through an exploration of various models including a condensed Grades K–8 growth model, a piecewise model with distinct Grades K–1 and Grades 3–8 slopes, and a model with unique Grades K–1 and Grades 3–8 mathematics achievement slopes and intercepts. The second aim was to assess the relationships between the mathematics achievement growth parameters and evaluate the extent to which mathematics growth trajectories differ in the early and middle grades. The third aim was to model Grades 3–8 mathematics self-efficacy growth trajectories and investigate the extent to which there was statistically significant growth in self-efficacy between Grades 3–8. The fourth aim was to investigate the relationship between mathematics achievement growth and the development of self-efficacy in mathematics to determine the unique relationships between Grades K–1 and Grades 3–8 mathematics achievement growth and mathematics self-efficacy development. The final aim was to evaluate the extent to which mathematics achievement growth, self-efficacy development, and the relationships between mathematics self-efficacy and mathematics achievement growth vary based on demographic factors such as sex, SES, and race/ethnicity. This final aim also included an investigation into the moderating relationships between mathematics achievement, self-efficacy, and demographic characteristics. These aims are captured by the following research questions:

1) Given a diverse, nationally representative sample of students, what is the most appropriate model for estimating mathematics achievement growth in Grades K–8?
2) To what extent is there evidence of unique mathematics achievement growth parameters in Grades K–1 and Grades 3–8 (i.e., are the slopes statistically significantly different in each time period)?
3) What are the relationships between Grades K–1 mathematics achievement growth parameters and Grades 3–8 mathematics achievement growth parameters?
4) To what extent are there significant growth parameters of mathematics self-efficacy in Grades 3–8?
5) What is the relationship between mathematics achievement growth and mathematics self-efficacy development in Grades 3–8 and does Grades 3–8 mathematics achievement growth mediate the relationship between Grades K–1 mathematics growth and Grades 3–8 mathematics self-efficacy development?
6) To what extent are there demographic (sex, SES, race/ethnicity) differences in mathematics achievement growth and mathematics self-efficacy development when controlling for other demographic factors?
7) How do demographic factors moderate the relationships between mathematics achievement growth and mathematics self-efficacy development?

CHAPTER III

METHOD

Data analyzed here were collected by the National Center for Educational Statistics (NCES) in the Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K). Academic and cognitive skills were measured, and demographic and self-report survey data were collected directly from children, their families, teachers, and schools over seven waves of data collection (Tourangeau, Nord, Le, Sorongon, & Najarian, 2009). Assessments and surveys were administered in the fall and spring of both kindergarten and first grade, and follow-up data collection occurred in the spring of third, fifth, and eighth grades.

Participants & Procedures

The ECLS-K study followed a nationally representative sample of 21,260 participating students from kindergarten to eighth grade beginning in the fall of 1998 through the 2007-08 school year. The study utilized a complex sampling design that included oversampling particular students (e.g., students from private schools and students from underrepresented races and ethnicities) to allow for various subgroup analyses. In all, the sample was drawn from 1,413 schools (953 public and 460 private) with 20,578 of 22,813 students responding by the spring of kindergarten and 8,706 children or 41% of the base-year respondents participating in all five years of data collection in Grades K–8 (Tourangeau et al., 2009).

Because attrition is expected in any longitudinal study, the ECLS-K sampling plan included a number of provisions to ensure that data collection was both feasible and fruitful. For example, in Grade 3 students who changed schools were sampled at a rate of

0.5 and students from language minority backgrounds were followed at 100%. Then, Grade 5 students whose home language was not English were still sampled at a higher rate, but children were subsampled at different rates depending on the longitudinal data available for each participant. Additionally, groups of students were not followed if they lacked data from previous rounds due to moving, being subsampled out, or parent refusal (see Tourangeau et al., 2009). This resulted in 5,214 students being excluded in Grade 5. By Grade 8, however, all eligible students (n = 12,129) were sampled regardless of status. The unweighted kindergarten and first grade ECLS-K sample was approximately 51% male, 52% White, 15% Black, 18% Hispanic, 6% Asian, and a combined 6% Native American, Alaskan Native, Pacific Islander, and/or more than one race, from a range of geographic areas, urban and rural schools, and a range of SES backgrounds.

Measures

At each data collection point, students were assessed using untimed, expert created, criterion-referenced, adaptive measures of cognitive ability including reading, mathematics, and general knowledge (science and social studies). To create the measures, expert consultants first identified target domains and constructs. Next, an item pool for each of the direct cognitive assessments was developed by adapting items from a variety of published assessments such as the Primary Test of Cognitive Skills (PTCS), the Test of Early Mathematics Ability (TEMA-2), and others. These items were then field tested for individual administration, with the exception of the Grade 8 measures, which were group administered. Finally, the resulting mathematics and general knowledge forms were translated to Spanish and administered to all students who were proficient in either English or Spanish in Grades K–1. Students whose primary language was something other than Spanish or English did not complete any of the direct cognitive assessments until their English language skills were proficient.

Because the ECLS-K assessments were supplementary to the participants' educational program, assessment developers aimed for a combined total test time of less than one hour for all of the direct cognitive assessments. This time constraint guided item development such that the direct cognitive assessments included 50-70 multiple choice and open-ended items. In the mathematics assessments item specifications and content distributions varied slightly each year; 40-50% of the items assessed number sense, properties, and operations, 15-20% assessed measurement, 5-15% assessed geometry and spatial sense, 10% assessed data analysis, statistics, and probability, and 15-20% assessed patterns, algebra, and functions (Rock & Pollack, 2002).

Time concerns also directed assessment administration. Students first completed a brief routing assessment to determine what level of second-stage assessment should be administered. Specific cut scores were developed through statistical simulations that employed IRT ability estimates and item difficulty parameters with the goal that test takers would be appropriately assigned to a high, medium, or low level assessment, as appropriate. The mathematics routing forms demonstrated an internal consistency alpha range between .76 and .88 across all assessment periods (Rock & Pollack, 2002; Tourangeau et al., 2009).

Item Response Theory (IRT) theta scores ranging from -3 to 3 were generated for each student in each subject area with higher scores representing better performance on the achievement measures. Because academic achievement was reported using IRT scoring, student scores in mathematics in the ECLS-K data set represent an estimation of each student's ability (theta) in mathematics at each measurement point. The resulting scores are reported on a vertically linked scale so scores can be compared over time and across measurement occasions. The ECLS-K manual reports IRT-derived assessment reliability ranging from .91 to .95 across each wave of data collection and percent agreement for the open-ended items ranged from .95 to .98 (Tourangeau et al., 2009). Validity evidence for the direct cognitive assessments is somewhat limited and based primarily on expert review and content mapping. The ECLS-K Grades K–1 psychometric manual notes that the mathematics assessments demonstrated a correlation with the other direct cognitive assessments (reading and general knowledge) ranging from .64 to .77 in those years (Rock & Pollack, 2002).

Demographic data collected in the fall of 1998 from school data systems, parent surveys, and student reports including sex, socioeconomic status, and race/ethnicity was utilized in the current study. Sex is a dichotomously coded variable representing male (coded as 1) and female (coded as 0). Socioeconomic status is represented by a continuous variable on the scale of -3.00 to 3.00 with larger values representing higher SES calculated as a composite from items collected in a parent survey including (a) male guardian's level of education, (b) female guardian's level of education, (c) male guardian's occupation, (d) female guardian's occupation, and (e) household income. Guardians' occupations were converted to prestige scores based on the General Social Survey of 1989 (Tourangeau et al., 2009). To facilitate demographic group comparisons, the socioeconomic status variable was dichotomized using a median split to generate above median (coded as 1) and below median (coded as 0) SES student groups. Race and ethnicity is a nominal variable containing the following categories: White (Non-

Hispanic), Black or African American (Non-Hispanic), Hispanic (Race Specified), Hispanic (Race Not Specified), Asian, Native Hawaiian (Other Pacific Islander), American Indian or Alaska Native, More Than One Race (Non-Hispanic). The race variable was dichotomized into two groups for group comparisons. The students from Black or African American, Hispanic (Race Specified), Hispanic (Race Not Specified), Native Hawaiian (Other Pacific Islander), American Indian or Alaska Native, and More Than One Race (Non-Hispanic) racial and ethnic backgrounds comprised the traditionally underserved group (coded as 1) and the White (Non-Hispanic) and Asian students were combined into a traditionally adequately served racial and ethnic group (coded as 0).

In the third, fifth, and eighth grades, NCES administered a student survey to measure intra and interpersonal skills, social and emotional factors, school adjustment, and health and wellness habits. Student survey questions of interest for the current study were adapted from the Self Description Questionnaire II (SDQ; Marsh, 1992) and included student ratings of agreement with the following statements pertaining to self-efficacy in mathematics: (a) Work in math is easy for me (Grades 3 and 5); (b) I can do very difficult problems in math (Grades 3 and 5); (c) I am good at math (Grades 3 and 5); (d) I get good grades in math (all grades); and (e) Math is one of my best subjects (Grade 8). Responses were reported on a 4-point Likert scale ranging from 1 = “not at all true” to 4 = “very true.” It should be noted that the ECLS-K dataset provides a Self-Description Questionnaire Perceived Interest/Competence in Math scale score that represents the mean rating of a number of mathematics-related survey items that go beyond self-efficacy; therefore, only the questions listed above were used to isolate self-efficacy in mathematics. Generating a combined raw score that represents student reported mathematics self-efficacy allows for an analysis of self-perceived competence or efficacy separate from interest in mathematics. This distinction is desirable based on research findings that suggest it is common for girls in middle and secondary grades to report lower levels of interest in or desire to do mathematics regardless of achievement (Frenzel, Pekrun, & Goetz, 2007) and competence beliefs are often closely intertwined with the perceived value of the domain (Fredricks & Eccles, 2002; Jacobs, Lanza, Osgood, Eccles, & Wigfield, 2002).

Data Analyses

Preliminary descriptive analyses were conducted using SPSS 20 (IBM Corp., 2011) and all subsequent models were investigated using robust maximum likelihood estimation with Mplus 7.1 (Muthén & Muthén, 2013). Univariate descriptive analyses were performed on measures of mathematics achievement, survey items, and student characteristics. Pearson's *r* correlation coefficients were used to examine the covariation among the study variables. Prior to conducting statistical analyses, all data was plotted, graphed, visually inspected to determine functional form and variable distributions (see Figure 1). Additionally, due to the large sample size utilized in this study and a desire to be conservative in any descriptive conclusions drawn from these analyses an alpha value of .01 was utilized for all analyses to identify substantive and educationally meaningful results.

To estimate the extent to which statistically significant mean growth parameters exist in Grades K–8 mathematics achievement, latent growth-curve modeling was employed to evaluate a set of nested mathematics growth trajectories spanning Grades K–8. For each mathematics achievement growth model, model fit was evaluated using various model fit statistics (Byrne, 2012; Kline, 2010). Specifically, two incremental indices including the chi-square goodness of fit test (H_0: The model fits the data) and the Comparative Fit Index (CFI; values of .95 or larger indicate acceptable fit); two predictive, parsimony-corrected fit indices including Aikaike's Information Criteria (AIC; smaller values are desirable) and Bayes Information Criteria (BIC: smaller values are desirable); and one absolute fit index-- Root Mean Square Error Absolute (RMSEA; values less than .08 indicate acceptable fit).

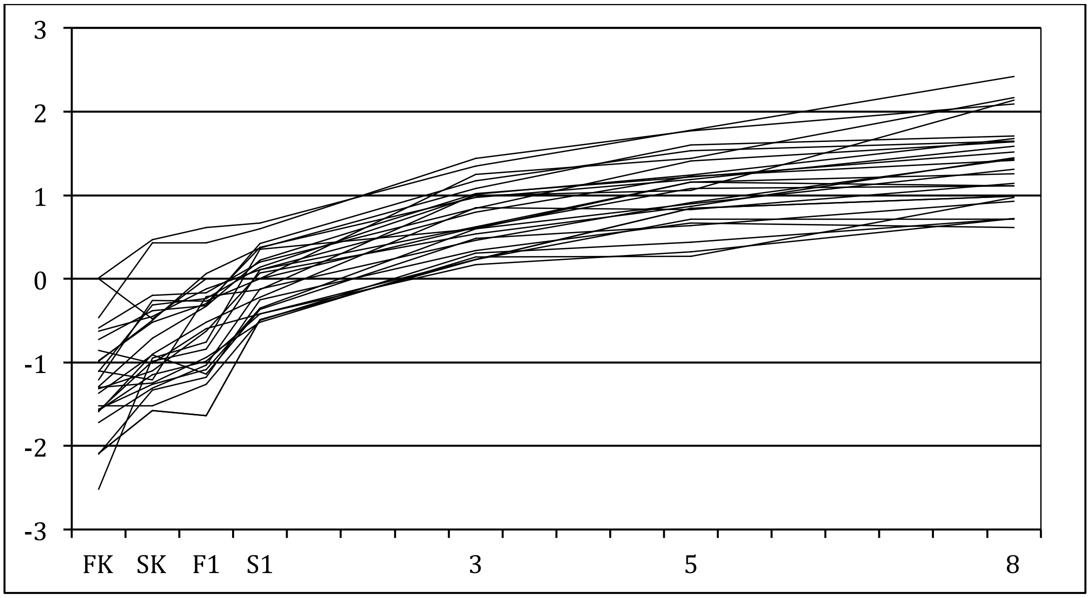

Figure 1. Random sample of 25 actual Grades K–8 mathematics achievement growth trajectories

Because the analyses described here were conducted utilizing robust maximum likelihood estimation, a scaling factor was provided for chi-square values and Satorra-Bentler corrected chi-square difference tests (Satorra & Bentler, 2001) were conducted for all models. The Satorra-Bentler corrected chi-square difference tests account for the Satorra-Bentler chi-square scaling correction factor, which is generated to adjust for any potential issues of non-normality and were utilized to evaluate whether there was evidence of statistically significant improvement in model fit.

First, a complete growth model for calendar year mathematics achievement growth across the kindergarten through eighth grade years was fit. A mean intercept and mean slope with time coded 0, .5, 1, 1.5, 3, 5, 8, as indicated in Table 1, was modeled from seven mathematics achievement measurements; two in kindergarten, two in first grade, one in third grade, one in fifth grade, and one in eighth grade. Next, a second complete Grades K–8 mathematics achievement growth model was generated with the slope term parameterized to represent academic year mathematics achievement growth (see Table 1). Because this model simply represented an alternate slope parameterization, AIC and BIC values were utilized to evaluate relative model fit.

After the most appropriate Grades K–8 mathematics slope parameterization was determined, an additional summer discontinuity parameter representing potential summer loss (see Cooper, Nye, Charlton, Lindsay, & Greathouse, 1996) between spring kindergarten and fall first grade was added to the third growth model, see Table 1for model parameterizations. The fit of this model was compared to Model 2 using nested model comparisons to determine if the summer discontinuity parameter should be retained.

Next, a quadratic slope was added in Model 4 based on previous research findings suggesting that the rate of academic growth often slows in the middle grades (Bloom, Hill, Black, & Lipsey, 2008) and preliminary visual analyses of actual growth trajectories. Because it was hypothesized that there would be substantively different mathematics achievement growth in the early grades as compared to the middle grades, and based on the plotted mean quadratic growth curve in Figure 2, another nested model comparison was conducted to examine whether a piecewise model could better represent mathematics achievement growth across Grades K–8.

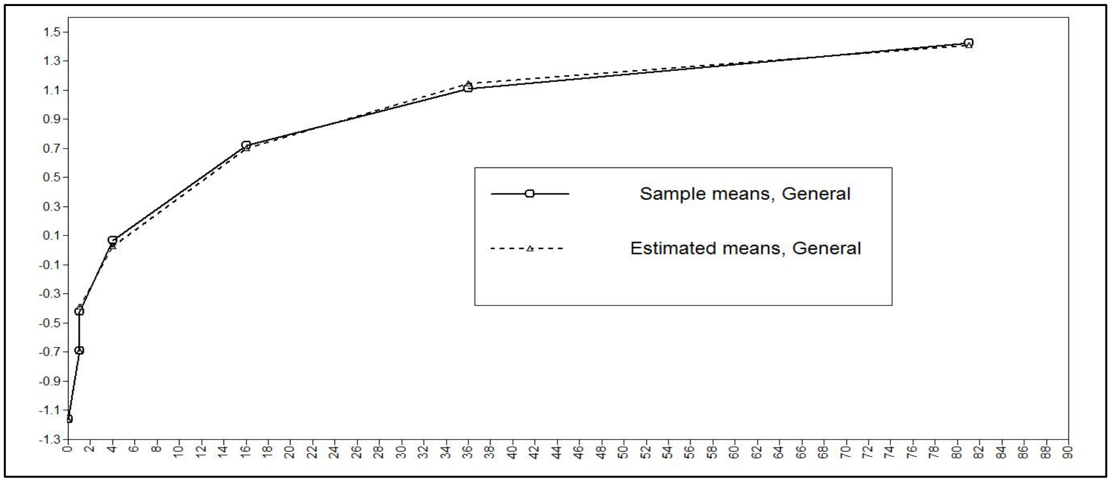

Figure 2. Estimated and actual mean Grades K–8 mathematics achievement growth trajectories

Model 5 (early grades + middle grades piecewise growth model) included separate Grades K–1 and Grades 3-8 cross-year slopes and intercepts with time coded as specified in Table 1. Lastly, two separate growth models reflecting Grades K–8 mathematics achievement growth in two distinct Grades K–1 and Grades 3–8 growth curves (each with its own unique intercept and slope) was generated in Model 6. Model 6 (split early grades & middle grades) was tested based on a desire to reflect (a) the differential nature and quality of achievement growth in Grades K–1 and Grades 3–8 (Bloom, Hill, Black, & Lipsey, 2008), (b) the measurement differences inherent in evaluating number sense development in the early grades as compared to assessments of concept integration and application in the middle grades, and (c) the practical utility of evaluating early and middle grades mathematics achievement growth separately for policy and instructional planning. See Table 1 for model parameterizations.

Table 1

Grades K–8 Mathematics Achievement Latent Growth Term Parameterizations

Model	Fall K	Spring K	Fall 1	Spring 1	3	5	8
Intercept (Models 1–5)	1	1	1	1	1	1	1
1. Calendar year slope	0	.5	1	1.5	3	5	8
2. Academic year slope	0	1	1	2	4	6	9
3. Academic year slope	0	1	1	2	4	6	9
Summer discontinuity	0	0	1	1	1	1	1
4. Academic year slope	0	1	1	2	4	6	9
Summer discontinuity	0	0	1	1	1	1	1
Quadratic slope	0	1	1	4	16	36	81
5. Early grades slope	0	1	1	2	2	2	2
Summer discontinuity	0	0	1	1	1	1	1
Middle grades intercept	0	0	0	0	1	1	1
Middle grades slope	0	0	0	0	0	2	5
6. Split early intercept	1	1	1	1	--	--	--
Split early slope	0	1	1	2	--	--	--
Split summer discontinuity	0	0	1	1	--	--	--
Split middle intercept	--	--	--	--	1	1	1
Split middle slope	--	--	--	--	0	2	5

Because Model 6 was not nested, relative fit was determined using AIC and BIC values in combination with other fit indices. The statistical model in Figure 3 reflects this anticipated final split growth model that utilizes continuous, longitudinal mathematics achievement measurements to generate intercepts and slopes (i.e., academic year and summer discontinuity parameter) and represents a more theoretical and parsimonious model of mathematics achievement growth.

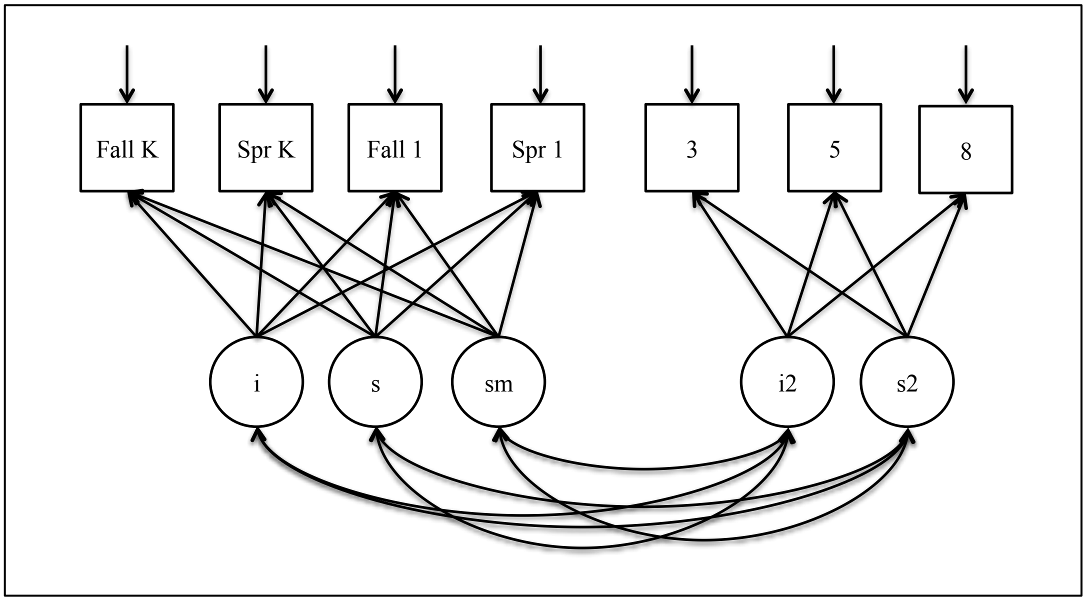

Figure 3. Split Grades K–1 and Grades 3–8 mathematics achievement growth model

After a final mathematics achievement growth model was established, relationships between the Grades K–1 growth parameters and Grades 3–8 growth parameters were investigated by evaluating the magnitude and statistical significance of the covariances between all growth terms. Additionally, to answer research question 2, the split mathematics achievement growth model was utilized to evaluate the extent to which the Grades K–1 cross-year slope was statistically significantly different from the Grades 3–8 cross-year slope. A Satorra-Bentler corrected chi-square difference test was conducted to compare a model where the two slopes were freely estimated to a model where the Grades K–1 and Grades 3–8 slopes were constrained to be equal.

Next, a mathematics self-efficacy growth model (see Figure 4) was developed using the mathematics self-efficacy mean scores produced by summing categorical response scores to student mathematics self-efficacy survey items from Grades 3–8. The mathematics self-efficacy slope was parameterized with time coded in the same manner as the Grades 3–8 mathematics achievement growth model (i.e., 0, 2, 5). Subsequently, to examine the relationships between mathematics achievement growth and mathematics self-efficacy development, the mathematics achievement and mathematics self-efficacy growth models were combined (see Figure 5) and mediation tests were conducted.

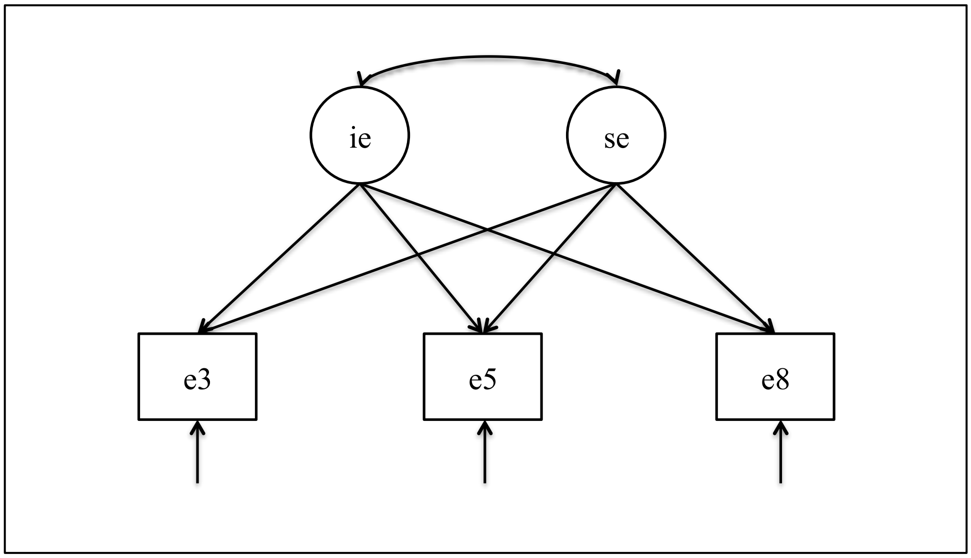

Figure 4. Grades 3–8 mathematics self-efficacy growth model

Direct and indirect effects were generated and evaluated using the following multiple step process:

1) The self-efficacy intercept and slope were regressed on the Grades K–1 intercept and slope.

2) The Grades K–1 intercept and slope were entered as predictors of the Grades 3–8 intercept and slope.
3) The mathematics self-efficacy growth parameters were regressed on the Grades 3–8 mathematics achievement intercept and slope.
4) Indirect effects were requested to examine the extent to which Grades 3–8 mathematics achievement growth mediates the relationship between mathematics self-efficacy development and early grades mathematics achievement.

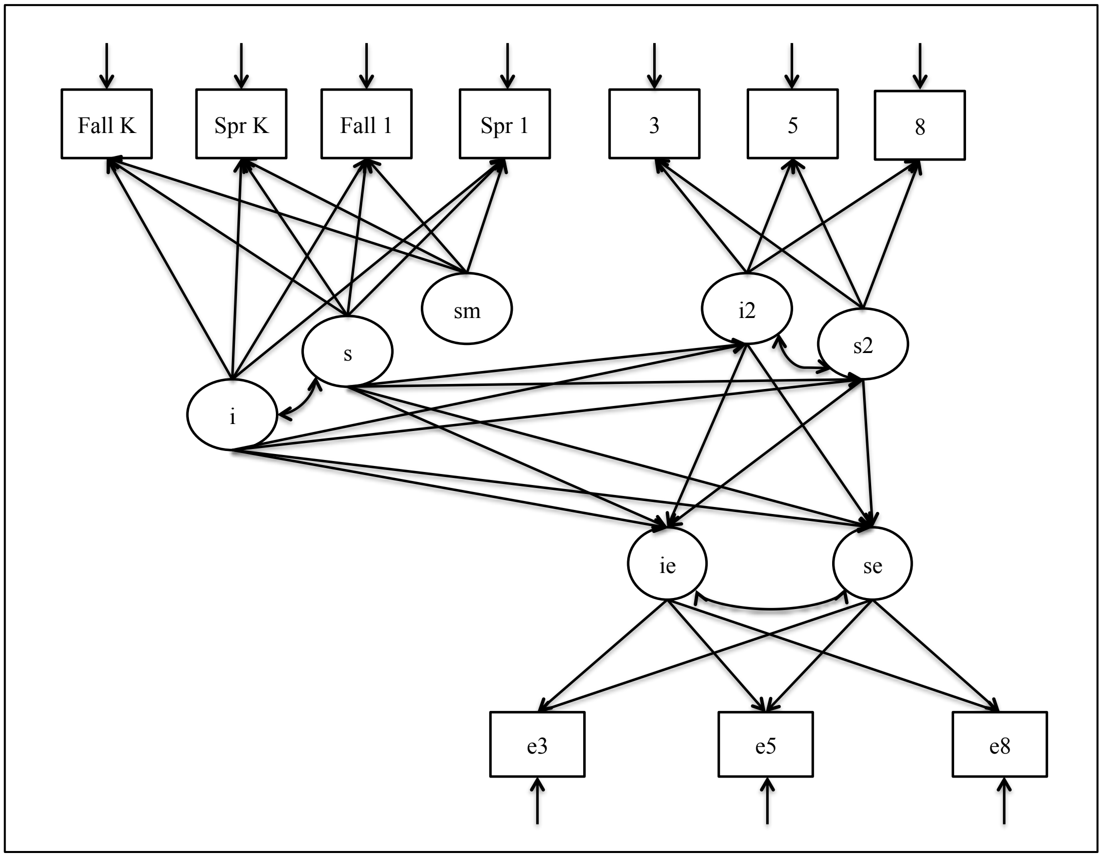

Figure 5. Grades K–1 and Grades 3–8 mathematics achievement growth and mathematics self-efficacy growth mediation testing model

The main effects of sex, SES, and race/ethnicity on mathematics achievement and mathematics self-efficacy development were also explored in a two-step process to allow for an examination of basic demographic effects followed by an investigation of the combined effects of mathematics achievement and demographic factors on self-efficacy development. First, all demographic factors were entered simultaneously and the growth parameters were allowed to covary freely to evaluate the extent to which mathematics achievement and mathematics self-efficacy growth parameters varied based on each demographic characteristic when controlling for the other demographic factors. Next, the covariances between mathematics achievement growth and self-efficacy development were replaced with the statistically significant predictive paths from the mediation model and the resulting model was used to identify all statistically significant associations between mathematics achievement growth, self-efficacy development, and demographic characteristics when controlling for all other factors.

Sex, SES, and race/ethnicity were also investigated as potential moderators of the relationships between the mathematics achievement growth and self-efficacy development. In this final analytic step, group difference tests were conducted to determine the extent to which sex, SES, and race/ethnicity moderated the relationships between mathematics achievement growth parameters and mathematics self-efficacy development trajectories. In each group difference test, the predictive paths between the mathematics achievement growth parameters and the mathematics self-efficacy growth parameters were first freely estimated and then constrained to be equal for each demographic group (i.e., males & females, above median SES & below median SES, and traditionally well-served racial/ethnic groups & traditionally underserved racial/ethnic groups). Satorra-Bentler corrected chi-square difference tests were conducted to determine the statistical significance of the change in model fit between each model. If initial chi-square tests revealed a statistically significant improvement in model fit when the paths were freed, follow-up tests were conducted to explore which specific model paths varied by sex, SES, and/or race/ethnicity.

Because the statistical computing software was unable to converge on a solution when the main effects of sex, SES, and race/ethnicity were freed in addition to the mathematics achievement and self-efficacy paths, separate group difference tests were conducted to investigate whether there was evidence of potential three-way interactions between a combination of demographic characteristics and the mathematics achievement and mathematics self-efficacy growth parameters. For these analyses, the predictive paths between the mathematics achievement growth parameters and the self-efficacy growth parameters were fixed and the main effects of the demographic factors were allowed to vary between groups. Satorra-Bentler corrected chi-square difference tests were utilized to examine the change in model fit between each model.

Sample weights. The ECLS-K dataset does not reflect a simple random sample. To ensure that data was collected from a wide range of children, particular schools, teachers, and children were targeted for participation, thus all subjects did not have an equal chance of being selected for the study. To account for selection and nonresponse bias resulting from the complex, stratified, multistage probability sampling design utilized in the ECLS-K study, a longitudinal child level sample weight will be applied to all analyses. Ignoring the unique structure of these data can result in the generation of inaccurate standard errors and misleading results (Stapleton, 2008). In contrast, applying the weighting variable will allow for the proposed analyses to most closely reflect a nationally representative sample.

Mplus allows for the application of a weighting variable, but the weight variable was adjusted to account for the sample design effects before being utilized in the analyses (Asparouhov, 2005). To generate a design effect (DEFF) adjusted weight variable, the weight variable was first normalized to reflect the analytic sample by multiplying the weight variable by a normalizing factor (i.e., the number of cases with nonzero weight values divided by the sum of the weight variable). Then, the normalized weight was divided by the appropriate DEFF value (3.043; Tourangeau et al., 2009). After applying the Grades K–8 longitudinal sample weight, the resulting sample included 2368 cases with non-zero weights.

Missing data. The ECLS-K study utilized a purposeful, stratified sampling plan and the dataset includes sampling weights that reflect this plan and account for sampling bias. However, cases that were missing all data utilized here (n = 11) and cases that were missing all mathematics achievement scores (n = 353) were omitted from these analyses prior to the application of the weighting variable. The extent to which there was systematic or relevant missingness was examined in a number of ways. Because cases that were missing mathematics theta scores from all waves of data collection were omitted from this analysis, missingness analyses were conducted to examine whether these cases were statistically significantly associated with the demographic variables. These analyses revealed that those missing mathematics achievement scores from all waves of data collection included fewer White students and more Asian students than expected ($\chi^2(7) = 101.96$, $p < .001$). Mathematics achievement score missingness was not statistically significantly associated with sex or reports of mathematics self-efficacy.

Although the robust maximum likelihood estimation utilized in these analyses enables statistical analyses to be conducted utilizing all available information from all participants, mathematics theta score missingness is reported here to determine the percentage of the sample that had (a) complete data across all seven time points, (b) at least one mathematics theta score in Grades K–1, (c) at least one mathematics theta score in Grades 3–8, and (d) at least one mathematics theta score in each analytic segment (i.e., Grades K–1 & Grades 3–8). See Tables 2 and 3 for mathematics achievement data percentages.

Table 2

*Missing Mathematics Theta Score Frequencies for Cases with Non-Zero Sample Weight (*n *= 2,368)*

Missing data points	*N*	%	Cum. %
0	2306	97.4	97.4
1	32	1.4	98.8
2	16	0.6	99.4
3	8	0.3	99.7
4	5	0.2	99.9
5	1	0.1	100.0
6	--	--	100.0

Note. Cases with missing data across all seven waves were omitted from analyses

Complete race and sex data was available for all participants with a non-zero sample weight (n = 2368), and 37 of these participants were missing SES data. Independent samples t-tests revealed that there were no statistically significant differences between the participants that were missing SES data and those that had an

SES value on measures of mathematics achievement or mathematics self-efficacy across all waves of data collection. There were also no statistically significant differences in sex, or race/ethnicity.

Table 3

*Missing Mathematics Theta Score Frequencies by Wave for Cases with Non-Zero Sample Weight (*n *= 2,368)*

Wave	*N*	%	Cum. %
1	45	1.9	1.9
2	23	1.0	2.9
3	14	0.6	3.5
4	7	0.3	3.8
5	4	0.2	4.0
6	9	0.4	4.4
7	12	0.5	4.9

Note. 99.9% of cases had at least one data point in waves 1–4 and 99.9% of cases had at least one data point in waves 5–7.

CHAPTER IV

RESULTS

Results from preliminary tests of modeling assumptions suggested that the data utilized in this study were robust to the assumption of normality without severe kurtosis or skew (i.e., skewness statistics ranged from -0.68 to 0.69). The data were also free of unduly influential outliers. Pearson's correlations revealed statistically significant correlations between most of the study variables (see Table 4), and an inspection of a random selection of 25 samples of mathematics achievement performance across Grades K–8 (see Figure 1 above) suggested the presence of potential negative quadratic form or differential growth in the early grades as compared to the middle grades.

Both initial Grades K–8 mathematics achievement growth models (i.e., Model 1: calendar year and Model 2: academic year) demonstrated poor model fit (see Table 5 for complete model fit statistics); however, Model 2 (academic year) demonstrated smaller AIC and BIC values and this slope parameterization was retained in future models. When the summer discontinuity parameter was added to Model 2 and freely estimated in Model 3 (academic year + summer discontinuity), it generated a negative variance that was not statistically significant. Therefore, the summer discontinuity variance was fixed at zero for all subsequent models to ensure proper model estimation. As a result, the covariances between the summer discontinuity term and other latent and exogenous variables were also fixed to zero and remained as such in future models. Given these modifications, the nested models 2 and 3 were compared and a Satorra-Bentler corrected chi-square difference test revealed that Model 3 (academic year + summer discontinuity) demonstrated statistically significantly better fit ($\Delta\chi^2(4) = 1268.02$, $p < .001$) as compared to Model 2 and this parameterization was retained for future nested model comparisons.

Model 4 (academic year + quadratic), which accounted for the variable rates of growth across Grades K–8 resulted in statistically significantly improved model fit, $\Delta\chi^2(5) = 2352.28$, $p < .001$. However, Model 5 (early grades + middle grades piecewise) did not result in statistically significantly improved model fit based on all model fit statistics. Finally, Model 6 (split early & middle) demonstrated improved relative fit over Model 4 with AIC values of 2464.03 and 2884.86, and BIC values of 2600.97 and 3009.34, respectively. Thus, Model 6 was retained for the remaining analyses. Model 6 explained over 80% of the variance in all of the mathematics achievement variables (see Table 6).

All estimated sample means were statistically significantly different from zero at $p < .001$. The mean Grades K–1 intercept was -1.17, the mean early grades slope was 0.48, and the mean summer discontinuity value was 0.26. In Grades 3–8, the mean intercept was 0.75 and the mean slope was 0.14. Results from the Grades K–1 and Grades 3–8 mathematics achievement slope invariance test indicated that the model in which the Grades K–1 and Grades 3–8 mean slope parameters were freely estimated, demonstrated statistically significantly better model fit, $\Delta\chi^2(1) = 457.52$, $p < .001$. Therefore, the rate of mathematics growth in Grades K–1 was statistically significantly different from the rate of mathematics achievement growth in Grades 3–8.

Table 4

*Pairwise Weighted Descriptive Statistics and Statistically Significant Correlations for All Study Variables (*n *= 3731)*

Variable	1	2	3	4	5	6	7	8	9	10	11	12	*% or M* (*SD*)
1. Sex: male													51.31%
2. SES: above median	--												49.89%
3. Race: underserved	--	-.33**											40.00%
4. Fall K math	--	.39**	-.33**										-1.16 (0.48)
5. Spr. K math	--	.39**	-.31**	.84**									-0.69 (0.46)
6. Fall G1 math	--	.37**	-.30**	.80**	.85**								-0.42 (0.46)
7. Spr. G1 math	--	.37**	-.28**	.72**	.77**	.81**							0.07 (0.40)
8. Grade 3 math	.16**	.38**	-.31**	.72**	.74**	.77**	.80**						0.72 (0.39)
9. Grade 5 math	.16**	.37**	-.30**	.68**	.70**	.73**	.77**	.88**					1.11 (0.41)
10. Grade 8 math	.11*	.39**	-.32**	.64**	.68**	.69**	.72**	.84**	.86**				1.42 (0.45)
11. Grade 3 efficacy	.20**	.07	--	.15**	.17**	.16**	.16**	.24**	.22**	.22**			3.15 (0.74)
12. Grade 5 efficacy	.20**	.13**	-.07	.26**	.28**	.30**	.32**	.36**	.39**	.39**	.41**		2.96 (0.80)
13. Grade 8 efficacy	.09	.17**	-.09	.26**	.30**	.26**	.26**	.30**	.33**	.39**	.24**	.39**	2.70 (0.92)

Note. Correlations calculated using pairwise deletion. All reported correlations are significant at $p < .05$ or better. *%* = percent of sample*, M* = mean, *SD* = standard deviation.

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

Table 5

Grades K–8 Mathematics Achievement Growth Model Fit

Model	χ^2	df	scale factor	AIC	BIC	RMSEA	CFI
1. Calendar year	4738.96	24	5.71	27942.72	28011.19	0.23	0.11
2. Academic year	4202.70	24	5.72	24925.61	24994.08	0.22	0.21
3. Academic year + summer	2735.97	20	5.41	15677.05	15770.41	0.19	0.48
4. Academic year + quadratic	369.10	15	5.40	2884.86	3009.34	0.08	0.93
5. Early grades + middle grades	456.18	14	5.12	3229.69	3360.41	0.09	0.92
6. Split early & middle	311.56	13	5.03	2464.03	2600.97	0.08	0.94

Note. χ^2 = chi-square statistic, RMSEA = Root Mean Square Error Absolute, CFI = Comparative Fit Index.

Table 6

Variance Explained in Endogenous Variables in Grades K–8 Mathematics Achievement Growth & Mathematics Self-Efficacy Development Models

	Split early & middle	Efficacy only	Mediation	Main effects
Mathematics achievement				
Fall kindergarten	.89	--	.89	.80
Spring kindergarten	.81	--	.81	.82
Fall first grade	.82	--	.81	.84
Spring first grade	.87	--	.87	.88
Third grade	.87	--	.91	.92
Fifth grade	.83	--	.82	.82
Eighth grade	.85	--	.89	.91
Mathematics self-efficacy				
Third grade	--	.41	.41	.35
Fifth grade	--	.41	.41	.39
Eighth grade	--	.59	.58	.56
Latent variables				
Grades K–1 intercept	--	--	--	.34
Grades K–1 slope	--	--	--	.09
Grades 3–8 intercept	--	--	.81	.78
Grades 3–8 slope	--	--	.02	.05
Self-efficacy intercept	--	--	.21	.38
Self-efficacy slope	--	--	.27	.34

Both of the early and middles grades mean intercepts demonstrated statistically significant variance (*var*(i) = 0.22 and *var*(i2) = 0.13), and the Grades K–1 mean slope also displayed statistically significant variance (*var*(s) = 0.02); however, the middle grades slope did not exhibit statistically significant variance. Whereas, the early grades mean intercept and mean slope parameters demonstrated statistically significant, negative covariance, *cov*(i,s) = -0.04, $p < .001$, the middle grades mean intercept and mean slope parameters exhibited a statistically significant, positive, but very small in magnitude covariance, *cov*(i2,s2) = 0.01, $p = .001$. Using the a priori alpha value of .01, the covariance between the Grades K–1 mean intercept and the Grades 3–8 mean intercept (*cov*(i,i2) = 0.14, $p < .001$) was the only statistically significant cross-model association

Despite the fact that it explained less of the variance in each of the efficacy items as compared to the mathematics items in the mathematics growth model (see Table 6), the mathematics self-efficacy growth model fit was quite good; $\chi^2(3) = 9.71$, $p = .021$, CFI = .96, and RMSEA = .03. The mean intercept was 3.15 and the mean slope was -0.09; both parameters were statistically significant, $p < .001$. Both growth parameters also demonstrated statistically significant variance, *var*(i) = 0.24 and *var*(s) = 0.02, $p < .001$. Lastly, the covariance between the mean self-efficacy intercept and the mean self-efficacy slope was not statistically significant.

The first step in the mediation analysis revealed that the Grades K–1 mean intercept was a statistically significant predictor of both the self-efficacy intercept and slope ($\beta_{ie\ i} = -0.40$ and $\beta_{se\ i} = 0.46$); however, the Grades K–1 slope was not statistically significantly related to either self-efficacy growth term. In contrast to the final mathematics growth model (Model 6) where the relationship between the Grades K–1 and Grades 3–8 intercepts was the only statistically significant at $p < .01$ association between growth parameters, the results of step 2 revealed that both the Grades K–1 intercept and Grades K–1 slope were statistically significant predictors of the Grades 3–8 intercept ($\beta_{i2\ i} = 1.10$ and $\beta_{i2\ s} = 0.52$). However, consistent with Model 6, neither Grades K–1 growth parameter was statistically significantly related to the Grades 3–8 mean slope. In the third step of mediation testing, the Grades 3–8 mean mathematics achievement intercept was statistically associated with the self-efficacy intercept, $\beta_{ie\ i2} = 0.73$, and the Grades 3–8 mean slope was also associated with the mean self-efficacy slope, $\beta_{se\ s2} = 0.46$. The Grades 3–8 mathematics intercept did not demonstrate a statistically significant relationship with the self-efficacy slope. See Figure 6 for a complete representation of all statistically significant covariances and standardized model paths.

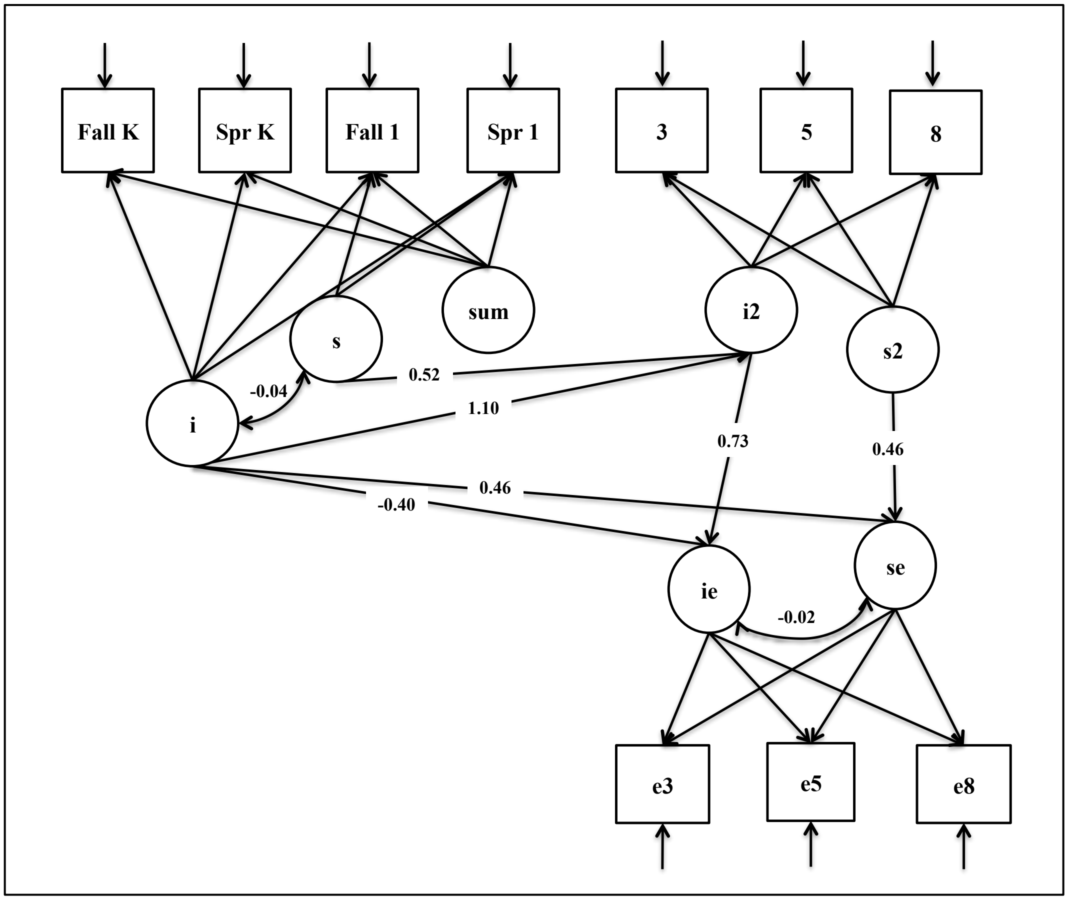

Figure 6. Statistically significant ($p < .01$) standardized beta weights and covariances describing the relationship between mathematics achievement growth and mathematics self-efficacy development in Grades K–8

The standardized indirect effects generated in the final step of mediation testing revealed a total indirect effect between the self-efficacy intercept and the Grades K–1 intercept via the Grades 3–8 intercept $\beta_{ie\ i2\ i} = 0.81$. This combined with the direct effect of $\beta_{\text{ie i}} = -0.40$ resulted in a total, positive relationship between the Grades K–1 intercept on the self-efficacy intercept of 0.41. There was also evidence of mediation in the relationship between the self-efficacy intercept and the Grades K–1 slope by the Grades 3–8 intercept with a resulting indirect effect of $\beta_{\text{ie i2 s}} = 0.38$. Finally, there was no evidence of mediation by the Grades 3–8 growth parameters in the relationship between the Grades K–1 growth parameters and the self-efficacy slope.

Group Differences

Results of the first demographic effects analysis revealed a number of main effects related to sex, SES, and race/ethnicity when the mathematics achievement self-efficacy growth parameters were allowed to freely covary. Specifically, the Grades K–1 intercept varied by SES and race/ethnicity such that students from traditionally underserved racial and ethnic backgrounds were predicted to have a Grades K–1 intercept that was 0.18 standard deviations lower than their traditionally well-served racial/ethnic peers ($\beta_{\text{i race}} = -0.18$, $p < .001$) and students from above median SES households were predicted to have a Grades K–1 intercept that was 0.44 standard deviations higher than their below median SES peers ($\beta_{\text{i ses}} = 0.44$, $p < .001$). The predicted rate of growth in Grades K–1 also varied by SES in that higher SES students were expected to demonstrate a slower rate of growth, $\beta_{\text{s ses}} = -0.19$, $p < .001$.

The Grades 3–8 intercept varied by SES ($\beta_{\text{i2 ses}} = 0.44$, $p < .001$), race/ethnicity ($\beta_{\text{i2 race}} = -0.17$, $p < .001$), and sex; male students were predicted to demonstrate a mean mathematics achievement score approximately 0.15 standard deviations higher than their female peers, $\beta_{\text{i2 sex}} = 0.15$, $p < .001$. Like the early grades, the Grades 3–8 slope also varied by SES, however in the middle grades students from higher SES households were predicted to demonstrate a rate of growth that was 0.25 standard deviations higher than their below median SES peers, $\beta_{s2\ ses} = 0.25$, $p = .007$.

There were also demographic differences in self-efficacy development. The level of self-reported mathematics self-efficacy in third grade differed by sex such that male students were predicted to report higher levels of self-efficacy, $\beta_{ie\ sex} = 0.33$, $p < .001$. Additionally, the slope of self-efficacy varied by SES with students from above median SES backgrounds expected to experience a positive improvement in the rate of self-efficacy growth between Grades 3–8 of 0.20 standard deviations, $\beta_{se\ ses} = 0.20$, $p = .003$. Although this is a substantively large positive effect, it did not change the valence of the self-efficacy slope term. The unstandardized effect for students who came from above median SES households was approximately 0.03, which resulted in a final predicted slope of (-0.04), so all students experienced decreasing levels of self-efficacy in the middle grades. See Figure 7 for a complete representation of all demographic main effects.

When the statistically significant regression paths between mathematics achievement growth and self-efficacy development (i.e., the paths in Figure 6) were replaced in the model, the relationships between early mathematics achievement growth and self-efficacy development disappeared and the magnitude of many of the relationships between Grades 3–8 mathematics achievement growth and self-efficacy development increased. The relationship between the Grades K–1 intercept and the Grades 3–8 intercept increased from 0.75 to 0.86 and the relationship between the Grades 3–8 mathematics achievement intercept and the self-efficacy intercept also increased from 0.45 to 0.59. The Grades 3–8 slope remained positively associated with the self-efficacy slope, $\beta_{se\ s2} = 0.53$, $p < .001$.

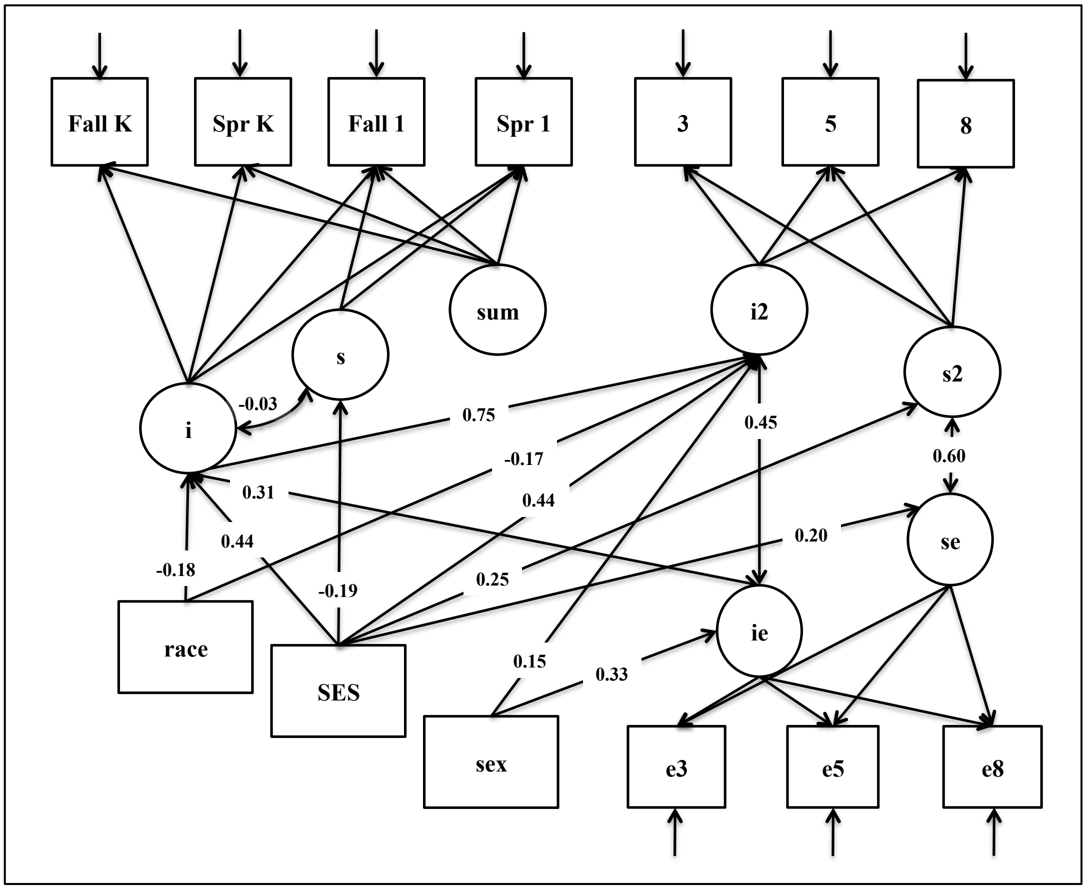

Figure 7. Statistically significant ($p < .01$) covariances between latent growth terms and standardized main effects of sex, SES, and race/ethnicity on Grades K–8 mathematics achievement and mathematics self-efficacy growth parameters

With the exception of sex, adding the regression paths to model also streamlined many of the demographic differences. Specifically, SES remained positively associated with both the Grades K–1 intercept and the Grades 3–8 slope, and negatively associated with Grades K–1 achievement growth, but was no longer associated with the Grades 3–8 intercept nor the self-efficacy slope. Likewise, race/ethnicity was only statistically significantly associated with the Grades K–1 intercept and there was no longer a direct relationship between race/ethnicity and differences in the Grades 3–8 mathematics achievement intercept. See Figure 8 for a complete model of the statistically significant relationships between mathematics achievement growth and self-efficacy development when controlling for the demographics factors.

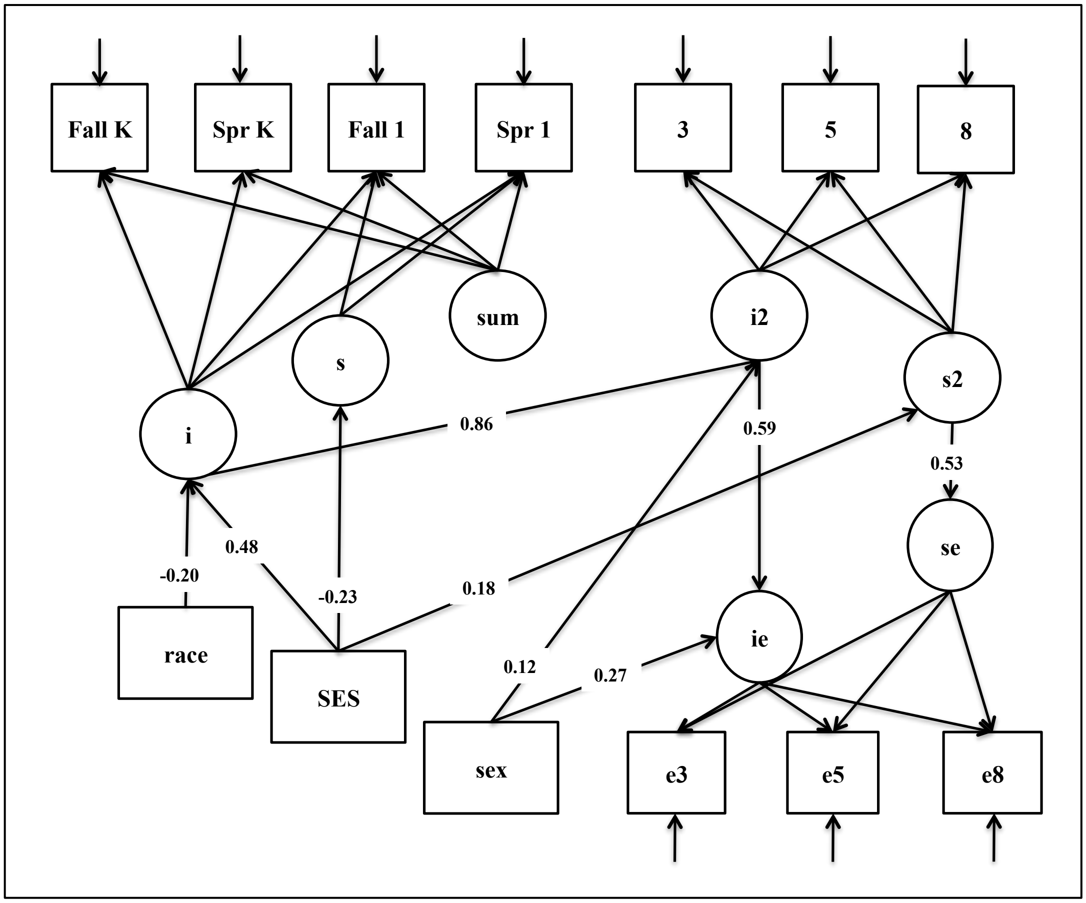

Figure 8. Statistically significant ($p < .01$) associations between Grades K–8 mathematics achievement and mathematics self-efficacy growth parameters when controlling for the standardized main effects of sex, SES, and race/ethnicity

Invariance testing was utilized to address the final research question and investigate the extent to which the relationships between mathematics achievement growth and self-efficacy development differed based on sex, SES, and/or race/ethnicity. Difference tests between the fixed and freed growth models for race/ethnicity and SES revealed no statistically significant differences in model fit. This result suggested that the relationships between mathematics achievement growth and mathematics self-efficacy development did not differ in a statistically significant manner based on race/ethnicity or

SES. Similarly, there was no evidence of three-way interactions between any combination of sex, SES, and race/ethnicity on either the mathematics achievement growth parameters or the mathematics self-efficacy growth parameters as none of these invariance tests were statistically significant.

The group difference test for male and female students did reveal a statistically significant difference in model fit for the freely estimated mathematics achievement and mathematics self-efficacy growth model (i.e., Model 6) as compared to the fixed model, $\Delta\chi^2(12) = 26.41$, $p = .009$. Follow-up parameter invariance tests revealed sex differences in the relationships between (a) the Grades K–1 intercept and the Grades 3–8 intercept, (b) the Grades K–1 intercept and the self-efficacy slope, (c) the Grades 3–8 intercept and the self-efficacy slope, and (d) the Grades 3–8 slope and the self-efficacy slope. See Figure 9 for male and female parameters with statistical significance at $p < .01$ indicated with an asterisk.

Although each of these relationships were statistically significantly different for male and female students, the Hedges' *g* effect sizes for each mean difference that consider both the sample size for each group and the standard error of each estimate were quite small (Hedges, 1981). For example, the effect size of the difference between the association between the Grades K–1 intercept and the Grades 3–8 intercept was $g = 0.01$. Similarly, the difference between the slope of self-efficacy development regressed on the Grades K–1 intercept and the Grades 3–8 intercept for male and female students corresponded to Hedges' *g* values of 0.01 and 0.04, respectively. Lastly, the difference in the relationship between the slope of self-efficacy development and the rate of Grades 3–8 mathematics achievement growth for boys and girls was $g = 0.04$.

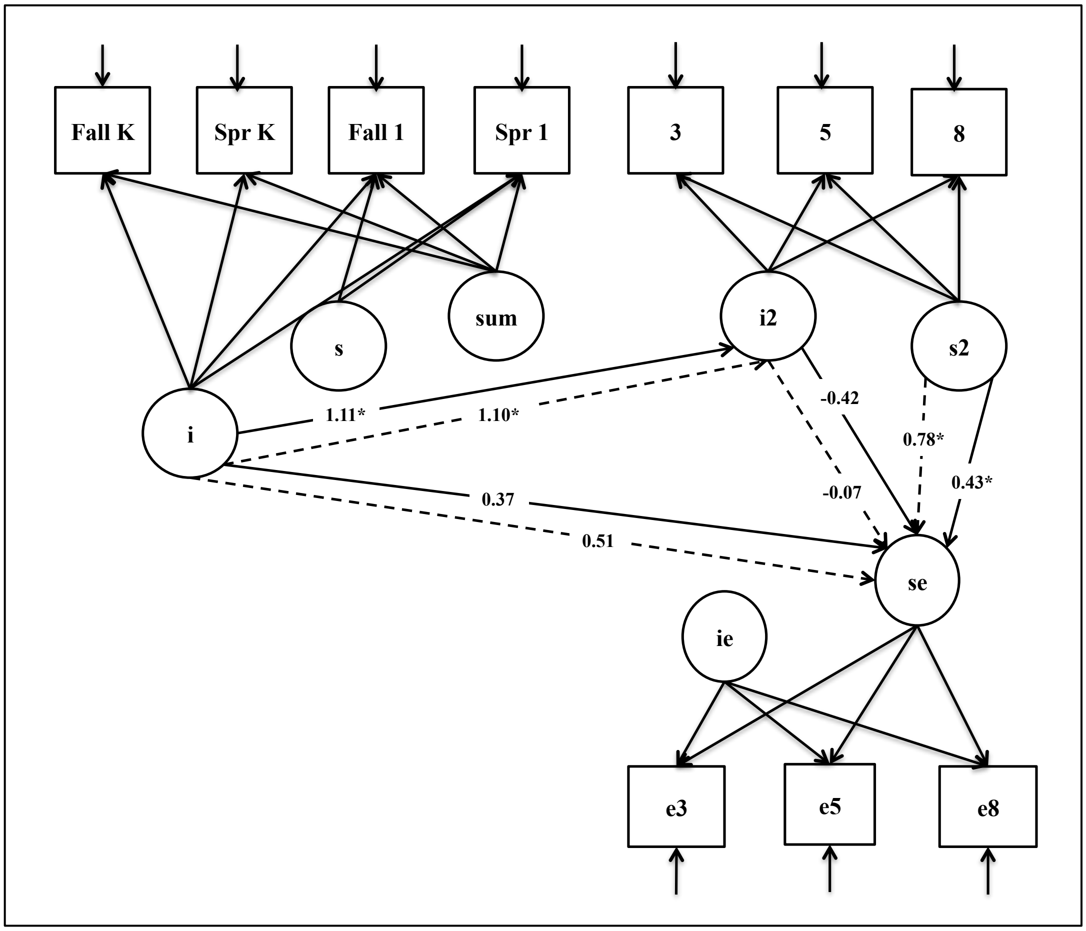

Figure 9. Statistically significantly different associations ($p < .01$) between mathematics achievement growth and the rate of development of mathematics self-efficacy for male (solid paths) and female (dashed paths) students

CHAPTER V

DISCUSSION

The current study yielded several findings regarding mathematics achievement growth and the development of self-efficacy in mathematics. While most findings were consistent with prior research, others provided novel contributions. The major findings are summarized briefly and then discussed in more detail below.

The superior fit of a split early and middle grades mathematics achievement growth model stands out as a unique and remarkable finding from the current study. As hypothesized, the rate of mathematics achievement growth in Grades 3–8 was significantly slower than the rate of mathematics achievement growth in kindergarten and first grade, and two separate growth models were best able to represent mathematics achievement growth across Grades K–8. Otherwise the features of mathematics achievement growth modeled in this study were largely consistent with previous research. Relationships between early and middle grades mathematics achievement revealed that early mathematics performance was highly correlated with later mathematics achievement. Additionally, there was a negative correlation between mathematics scores at kindergarten entry and subsequent rates of mathematics learning in kindergarten and first grade suggesting that the less students knew coming in, the more they had to learn and the quicker they acquired that basic knowledge. In contrast, the correlation between scores in third grade and rates of learning over the intermediate and middle school years was positive indicating that students who had higher mathematics scores in third grade also experienced improved rates of growth across the middle grades.

The development of self-efficacy also behaved as hypothesized with students reporting declining levels of self-efficacy across the middle grades. Mathematics achievement at kindergarten entry was uniquely associated with third grade reports of self-efficacy and the development of self-efficacy through eighth grade. In addition, mathematic achievement scores in third grade largely mediated the relationships between early mathematics achievement growth—both achievement at entry and growth over kindergarten and first grade—and third grade reports of self-efficacy.

Finally, all aspects of mathematics achievement growth and the development of self-efficacy were associated with at least one demographic factor. Generally these differences favored male students, White or Asian students, and students from above median SES households, and race/ethnicity and SES differences in self-efficacy development appeared to be moderated by mathematics achievement. The relationships between mathematics achievement and self-efficacy development sometimes reflected small but important differences between males and females. However, there was no evidence of differences in the relationships between mathematics achievement and self-efficacy development based on SES or race/ethnicity.

Split Model of Mathematics Growth

Consistent with previous research (see Ding & Davison, 2005; Cutuli et al., 2013), initial visual and statistical analyses suggested that mathematics achievement growth between kindergarten and eighth grade was not linear. Although a quadratic growth model with a constant linear slope and a quadratic slope that accounted for the reduced rates of growth in the middle grades fit the data, representing mathematics achievement growth in two separate growth models (Grades K–1 & Grades 3–8) with unique average initial performance scores and different rates of growth proved to be an even better fit. Quantifying the change in mathematics achievement over Grades K-8 with two distinct rates of growth leads to some interesting implications.

The fact that students demonstrated a much faster average rate of growth in kindergarten and first grades as compared to Grades 3–8 could be a result of a number of known and unknown factors. For example, it could be that students simply learn at a faster rate in the early grades because they have young, flexible minds and a generally good attitude toward school, or because there is just more for them to learn as they know so little at kindergarten entry. However, it could also be that the sensitivity of early mathematics assessments, which test discrete whole number skills like counting, cardinality, number identification, and basic facts (e.g., Test of Early Mathematics, Ginsburg & Baroody, 1983) lend themselves to more impressive rates of growth, whereas the more diversified and applied nature of intermediate and middle grades mathematics knowledge makes it more difficult to track and measure progress or growth. Quickly, easily, and efficiently assessing Grades 3–8 mathematics presents great challenges with respect to testing logistics and domain coverage. As the domain of mathematics becomes larger throughout these grades, the accumulation of skills, applications, and levels of knowledge to be assessed becomes greater and greater. However, time constraints and item formats that allow for expeditious scoring sometimes limit the depth and breadth of these assessments, and may limit the extent to which scores on these middle grades mathematics measures are valid approximations of the full breadth and depth of mathematics achievement (Reckase, 2004). For example, mathematics tests in Grades K–1 measure whole number skills like number identification, quantity discrimination, and basic number combinations, while tests in Grade 3 focus primarily on arithmetic skills, and eighth grade mathematics tests require the application of these early skills to problem solving, basic geometry, and algebraic reasoning. This would be akin to measuring reading growth using a combination of letter identification, phoneme segmentation, and word reading fluency in Grades K–1, oral reading fluency in Grade 3, and passage reading comprehension in eighth grade. Consequently, vertical comparisons of these substantively different constructs of mathematics knowledge may not be the most appropriate evaluation of mathematics achievement over time. Ergo, two separate models of growth were most appropriate.

One key finding from the split models was that although the average rates of change were statistically different between the early and middle grades, both of the mathematics slope parameters had little to no significant variance. This phenomenon could have a number of potential explanations. First, it may be that students simply accumulate mathematics knowledge at the same rate during both periods. Alternatively, it may be that the mathematics achievement measures utilized and/or the constructs assessed in this study were not sensitive enough to detect variations in mathematics achievement growth in either period. In contrast to the current findings, in at least one recent study the measurements of growth gathered using triannual benchmark assessments in middle school revealed significant variations in mathematics growth within a school year (Keller-Margulis, Mercer, & Shapiro, 2012). Similarly, Jordan and colleagues (2003) conducted a longitudinal analysis of the variations in rates of mathematics problem solving performance growth for elementary-aged students with high and low arithmetic fact mastery and found significant variance in growth for these targeted skills. Therefore, more sensitive assessments, such as curriculum based measures or progress monitoring tools validated for use with middle school students (Foegen, 2008) and more fine-grained analyses of specific mathematics skills across all grade levels might have revealed variance in growth that the measures in the current study could not detect. Finally, it could also be that one *could* see variance in mathematics achievement growth, even with the measures the current study used, if we were able to intervene more effectively with low performing students. In other words, students with less mathematics knowledge might be able to learn more quickly or “catch up” with their more advanced peers.

It is also noteworthy that although mathematics achievement scores at kindergarten entry were closely related to mathematics achievement in Grade 3 (as hypothesized; see Claessens, Duncan, & Engel, 2009; Duncan & Murnane, 2011), they were not associated with the rate at which students learn mathematics in Grades 3–8. Moreover, the rate at which students acquire mathematics in kindergarten and first grade was not highly related to mathematics achievement in third grade or subsequent growth. However, as noted above, there was very little variance present in rates of growth in the models examined here. Thus, it could be that there were not statistically significant relationships amongst the slope parameters due to this lack of variance. In fact, if we were able to (a) employ more sensitive measures of middle grades mathematics, or (b) develop timely and effective middle grades mathematics interventions to enable students with less knowledge to catch up to their average achieving peers more quickly thereby increasing the variance in rates of mathematics achievement growth, we might find different relationships between early and middle grades mathematics achievement growth.

Summer lag. Although the relationships between the Grades K–1 summer parameter and the other growth terms were not quantified here, the mathematics growth model building process in this study did confirm the presence of a statistically significant summer discontinuity term. This finding was consistent with research suggesting the presence of a summer slide where students experience reduced rates of growth or often achievement loss in the summer months (Cooper et al., 1996).

Although important to the stability of the overall model, the summer discontinuity term was somewhat problematic. In order for the mathematics achievement model to be properly estimated, the variance was fixed to zero suggesting that all students experienced a very similar, small mean growth rate between kindergarten and first grade. The positive summer discontinuity term generated in this study was somewhat unexpected given previous research on summer learning (Alexander, Entwisle, & Olson, 2001; Downey, von Hippel, & Hughes, 2008). Nonetheless, the current finding of small but significant growth during the summer was consistent with an earlier study of summer learning between Grades K and 1 using the ECLS-K dataset that reported very similar average summer gains (Burkam, Ready, Lee, & LoGerfo, 2004). Despite the fact that this finding was consistent with another study using the same dataset, these results should be interpreted with caution for the following reasons. For one, the fall first grade sample was much smaller than all other waves and was primarily intended to refresh the sample and capture students who did not attend kindergarten. Secondly, because evaluating summer loss was not the primary intent of the ECLS-K study, the actual date of data collection in each wave varied largely across the entire sample. As a result, the overall elapsed calendar time and amount of school-based instruction delivered between the spring of kindergarten and the fall of first grade was sometimes quite extensive. Next, because the statistical model required that the variance of the summer slope was fixed at zero, the associations between the summer discontinuity term and other mathematics achievement and self-efficacy terms remain unknown. Lastly, although previous research found demographic differences in summer learning (Burkam et al., 2004; Kim, 2004), the extent to which summer mathematics achievement differs for students based on sex, SES, and race/ethnicity could not be examined here due to the variance constraint.

Mathematics Self-Efficacy Development

Students reported a relatively high initial level of self-efficacy—3.15 on a scale from 1–4—suggesting that third graders in this nationally representative sample had generally high average levels of mathematics self-efficacy. However, there was statistically significant variance in these initial levels of mathematics self-efficacy and the standard deviation of the average self-efficacy score was approximately 0.75, so approximately 70% of the participating students reported average levels of self-efficacy between 2.40 and 3.90. It should be emphasized though that there is no absolute criterion for high self-efficacy or low self-efficacy on this scale. In fact, less than 5% of all sampled students reported responses that corresponded with a score of 1 on each question, so a score of 3 might actually reflect only moderate amounts of self-efficacy. Additionally, as noted in the introduction, assessments of self-efficacy are fraught with measurement challenges (Pajares, Miller, & Johnson, 1999; Wigfield, Eccles, & Pintrich, 1996) and speculations about scores on this scale should be made with caution.

The negative rate of self-efficacy development, or tendency to report lower levels of mathematics self-efficacy in the middle school years was consistent with previous research findings (Pintrich & Schunk, 1996) and suggested that students felt progressively less efficacious or less confident in their mathematics abilities in fifth and eighth grade as compared to third grade. These findings could be explained by the teenage developmental inclinations whereby middle school students tend to demonstrate less engagement and less motivation as compared to their younger peers (Blum & Libbey, 2004). However, reduced rates of mathematics self-efficacy could also be a real reflection of the development of negative mathematics identities commensurate with reduced rates of mathematics growth in the middle grades and the increasing complexity of mathematics content in the middle grades.

Relationships Between Mathematics Achievement and Self-Efficacy

Students with higher initial levels of mathematics achievement at school entry were expected to (a) have higher levels of mathematics achievement in third grade, and (b) experience less loss in self-efficacy development across Grades 3–8. Likewise, students who experienced better rates of mathematics achievement growth in the middle grades were expected to experience more positive development of self-reported self-efficacy across the middle grades. Specifically, reports of self-efficacy were expected to (a) hold fairly steady for students who demonstrated average mathematics growth in the middle grades, (b) decrease for students who experienced slower than average mathematics growth, and (c) increase for students who acquired mathematics more quickly than their peers. This relationship, combined with the fact that there was no statistically significant relationship between the rate of growth in Grades K–1 and the rate of self-efficacy development, provides further support for the importance of middle grades mathematics instruction. This relationship between Grades 3–8 mathematics growth and concurrent self-efficacy development is consistent with the theorized reciprocal nature of the relationship between academic achievement and self-efficacy (Caprara et al., 2011; Diseth, 2011). In short, effective middle grades mathematics interventions that improve rates of mathematics achievement can potentially enhance self-efficacy development, as well.

In addition to these direct relationships, mathematics achievement in third grade proved to be an important mediator of early mathematics and self-efficacy. Specifically, students with higher levels of kindergarten mathematics performance were predicted to have higher mathematics achievement scores in third grade and report higher levels of mathematics self-efficacy in third grade. Additionally, students who experienced better rates of mathematics achievement growth in kindergarten and first grade were also expected to have higher level of mathematics achievement in third grade and thereby report higher levels of self-efficacy. These findings suggest that although proximal and concurrent mathematics achievement is important to the development of mathematics self-efficacy, early experiences are also important. Early intervention programs that increase mathematics knowledge and student fluency with basic whole number concepts may in fact impact the long-term development of mathematics self-efficacy. Furthermore, students with positive early mathematics achievement experiences (both at kindergarten entry and across kindergarten and first grade) may be more inclined to report feeling efficacious in mathematics many years later. For example, interviews of eighth grade students with low mathematics self-efficacy revealed that both the students and their parents often attributed their low self-efficacy and difficulties in mathematics to shaky mathematics foundations and poor instruction in elementary school (Usher, 2009). This suggests that early academic experiences contribute not only to future achievement, but also the development of self-efficacy.

Group Differences in Mathematics Achievement and Self-Efficacy Growth

As expected, when specific relationships between the mathematics achievement growth and self-efficacy development were not specified, many aspects of mathematics achievement growth and self-efficacy development differed based on sex, SES, and race/ethnicity. All aspects of mathematics achievement growth and self-efficacy development demonstrated differences based on one or more of the demographic variables, but no one demographic was associated with differences in all aspects of mathematics achievement growth and self-efficacy development. Therefore, results are discussed on terms of each demographic characteristic and the aspects of mathematics achievement growth and self-efficacy development with which it was associated.

Race and ethnicity. Consistent with recent achievement gap literature (Reardon, Robinson, & Weathers, 2014), students from traditionally underserved racial backgrounds demonstrated lower levels of mathematics achievement both at kindergarten entry and in third grade. In fact, the magnitude of the difference between White and Asian students and other students of color was nearly identical in kindergarten and third grade, suggesting that there was little to no narrowing of the racial achievement gap in the early grades. Other examinations of various longitudinal datasets have found that most of the growth of the race/ethnicity achievement gap occurs in the early grades with the gap often widening across the elementary grades (see Murnane, Willett, Bub, &

McCartney, 2006; Reardon, Robinson, & Weathers, 2014). However, race/ethnicity differences were examined when controlling for both sex and SES in this study, and this may explain these more modest, yet important findings.

SES. Overall, SES-based differences were much more prevalent than race/ethnicity differences and revealed a similar picture of the mathematics achievement gap not closing appreciably. SES differences favored high SES children with one notable exception. The rate of mathematics achievement growth in Grades K–1 was predicted to be slightly slower for students from above median SES households than for those from below median households. Although somewhat counterintuitive, this common phenomenon often occurs as a result of students with higher initial scores experiencing somewhat slower rates of growth because they already have mastery of foundational knowledge and concepts, and acquiring these basic skills improves the rates of growth for students with lower initial achievement scores. It should also be noted that even though this difference was statistically significant, it was quite small, and because student from above median SES households demonstrated higher average scores at kindergarten entry, their scores remained higher than their lower SES peers across Grades K–1 in spite of their slower rate of growth.

Beyond the K–1 period, SES was also the only significant factor that impacted both the rate of mathematics achievement growth and the rate of self-efficacy development in Grades 3–8. In fact, higher SES students were predicted to have higher rates of mathematics learning in the middle years and were also expected to demonstrate more advantageous rates of self-efficacy development. Given that there was little variation in Grades 3–8 mathematics growth, the standardized effect for SES on mathematics growth was substantial; however, it was practically quite small with below median SES students expected to gain approximately 0.15 points per year and above median SES students predicted to gain 0.16. In contrast, although the standardized effect of SES on self-efficacy development was smaller than the effect on mathematics growth, the practical implications were more noticeable. Specifically, students from above median SES households were predicted to report decreases in self-efficacy at a rate that was half as fast as their peers from below median SES households. When combined with the predicted increase in self-efficacy development resulting from improved rates of mathematics growth in Grades 3–8, students from higher SES backgrounds demonstrated gains in both achievement and self-efficacy. Based on observed correlations between SES and both achievement and self-efficacy (Schunk & Meece, 2006; Sirin, 2005), these findings are consistent with SES-based achievement gap research (Reardon, 2011) and confirmed study hypotheses.

Sex. Sex differences did not emerge until Grade 3 at which point male students demonstrated higher mathematics achievement scores than did female students. The later emergence of sex differences in mathematics achievement is consistent with previous research findings (e.g., Hyde, Fennema, Ryan, Frost, & Hopp, 1990; Robinson & Lubienski, 2011) and may support previous research conducted with college-aged students suggesting that many female students learn to internalize gender stereotypes about achievement in mathematics and progressively view mathematics as a male domain (see Nosek, Banaji, & Greenwald, 2002), both of which negatively impact their mathematics achievement (Schmader, Johns, Barquissau, 2004). However, it should be noted that recent research exploring the role of stereotype threat in longitudinal mathematics achievement primarily conducted with students in the middle grades has failed to identify firm indicators of both (a) the emergence of stereotype threat in female students and (b) a clear relationship between expressed stereotype threat and poorer mathematics achievement (Ganley et al., 2013). Furthermore, results from a recent survey of almost 400 fifth grade students suggested that students in the middle grades have firm stereotypical beliefs about mathematics favoring males when polled about adult abilities, but beliefs about adolescent males being better at mathematics were less pronounced (Martinot, Bagès, & Désert, 2012). Of course, it is also possible that acquired mathematics anxiety could explain later sex-based differences in mathematics achievement, as mathematics anxiety has been shown to potentially affect the mathematics performance of females more than males (Devine, Fawcett, Szucs, & Dowker, 2012).

Social experiences and subconscious psychological constructs like fixed mindsets (Dweck, 2008) may also help to explain the fact that male students reported higher levels of mathematics self-efficacy in third grade, as well. Although the causal mechanisms remain somewhat unclear, there is also a good deal of research that suggests that teachers (and parents) possess, and may transfer beliefs about sex-based differences in mathematics achievement (Gunderson, Ramirez, Levine, Beilock, 2012; Keller, 2001; Li, 1999). For example, investigations of elementary school teachers' attitudes toward mathematics achievement and learning have revealed various sex differences. Teachers perceived male students as possessing more innate mathematics ability and believed that female students needed to be more effortful to achieve the same as their male peers in mathematics (Fennema, Peterson, Carpenter, & Lubienski, 1990). Similarly, interviews of eighth grade students revealed that (a) males with high mathematics self-efficacy often spoke of their mathematics achievement as the result of innate mathematics ability, and (b) students with low mathematics self-efficacy often reported that their mothers (and sometimes both parents) struggled with mathematics in school (Usher, 2009). The diverse findings of studies of sex differences in mathematics achievement and self-efficacy suggest that there are a variety of social, contextual, and psychological factors at play.

Controlling for achievement and efficacy relationships. Adding the predictive paths between mathematics achievement and self-efficacy development resulted in two important changes. First, the direct relationship between mathematics scores at kindergarten entry and reports of self-efficacy in third grade was not significant when the demographic factors were included. Therefore, when controlling for demographic characteristics and the relationship between kindergarten mathematics achievement and third grade mathematics scores, there was no independent relationship between kindergarten mathematics scores and reports of self-efficacy in third grade. This suggests that instructional efforts from kindergarten to third grade that improve third grade mathematics scores can have a meaningful effect on both future mathematics achievement and reports of self-efficacy in third grade.

Secondly, many of the demographic effects were reduced when directional relationships between mathematics achievement and self-efficacy development were specified in the model. While sex differences favoring males in mathematics achievement and reports of self-efficacy in third grade remained the same, differences based on SES and race/ethnicity changed. Mediation testing was not an explicit aim of this model, but the results suggest that the effects of both SES and race/ethnicity on mathematics achievement in Grade 3 was mediated by mathematics achievement at kindergarten entry and the effect of SES on the rate of self-efficacy development was similarly mediated by the rate of mathematics growth in the middle grades. In short, SES and race/ethnicity differences in third grade mathematics scores depended on the close relationship between kindergarten mathematics scores and third grade achievement. Similarly, the effect of SES on the rate of self-efficacy development appeared to work in concert with the relationship between Grades 3–8 mathematics growth and self-efficacy development. These findings were consistent with those discussed in the previous section with the effects of race/ethnicity and SES on mathematics achievement remaining largely constant in the early years and SES being associated with the rate of mathematics growth in Grades 3–8 and thereby the rate of self-efficacy development in the same time period. The fact that sex differences were unchanged suggested that there might also be sex differences in the *relationships between* mathematics achievement and self-efficacy.

Moderation of achievement and efficacy relationships. Whereas mathematics achievement scores, levels of self-efficacy, and the rates of growth for both differed based on demographic characteristics, the only significant differences in the relationships between mathematics achievement growth and the development of self-efficacy were based on sex. Although the differences were quite small—effects sizes ranged from .01–.04—the presence of significantly different estimates of the relationships between mathematics achievement growth and self-efficacy development for males and females, particularly when there were no SES or race/ethnicity differences suggests that the relationships between mathematics achievement and mathematics self-efficacy may differ based on sex.

Consequently, supporting mathematics achievement and continued growth may be especially important for female students as they may reap additional improvements in self-efficacy as a result. Specifically, mathematics achievement at kindergarten entry and mathematics achievement growth in Grades 3–8 were more strongly positively related to improved self-efficacy development for females, as opposed to males. These findings suggest that fostering strong foundational understandings of mathematics as early as kindergarten may support the development of feelings of efficacy toward mathematics in Grades 3–8. Additionally, supporting continued positive mathematics achievement growth in the middle grades may be especially critical to promote continued mathematics self-efficacy development, especially for female students. In fact, a recent study of American student performance on the Program of International Student Assessment (PISA) found that sex-based differences in mathematics achievement disappeared when controlling for mathematics self-efficacy (Kitsantas, Cheema, & Ware, 2011), suggesting that improved self-efficacy could help bridge achievement gaps. Because success and self-efficacy go hand-in-hand, targeted interventions should not only aim to convince students that they have the ability achieve, but also provide structured mathematics activities appropriate to each student's level so that they can experience a high level of success (Bandura, 1986). Interventions could also remind instructors to use frequent, carefully structured praise that focuses on effort and perseverance, rather than intelligence or innate ability (Usher, 2009). Providing structured opportunities for all students to leverage their mathematics skills to build mathematics efficacy thereby improving intra and interpersonal skills in addition to mathematics achievement, may be effective in bridging the gaps between male and female mathematics achievement and future outcomes.

Limitations & Future Directions

The findings from the current study must be tempered by acknowledging the limitations of the study, which were many. For one, the current study made use of an unusually large and nationally representative dataset, which offered extraordinary statistical power and generalizability. Nonetheless, the use of an extant, observational dataset limits the available measures and indicators for each construct of interest. There may be additional survey items or mathematics assessment tasks that would more comprehensively represent mathematics self-efficacy and achievement, and this is an important limitation to the present study. In addition to the limited ability of these data to represent the constructs of interest, there are a number of additional relevant covariates that are available in this dataset that could be incorporated in the analyses described here including school type, retention status, age at school entry, instructional variables, and even reading achievement. Furthermore, although predictive paths were modeled in these analyses, this study is purely descriptive and intended to describe and explore potential relationships between mathematics achievement, self-efficacy, and demographic characteristics.

Despite the strengths of the ECLS-K design and the generalizability it affords, there are also data collection issues that limit the generalizations that can be drawn from these analyses. For example, after kindergarten, data collection occurred whether or not the participating student was actually in the target grade, so if a student was retained or somehow skipped a grade, he was still assessed and the resulting data were included with the majority grade year. In the Grade 8 sample collected during the 2006-07 school year, approximately 89% of the students were in eighth grade classrooms with approximately 9% of the students in seventh grade and just over 1% in some other grade (e.g., tenth, sixth, or fifth; Tourangeau et al., 2009). Additionally, as noted earlier the ECLS-K sample was refreshed in first grade to include students who did not attend kindergarten. Thus, results from these analyses can only be generalized to US students who were in kindergarten in 1998-99 and/or first grade in 1999-2000 (Tourangeau et al., 2009).

Beyond limitations of the data, the choice of analytical approach also results in both strengths and limitations. General latent growth curve modeling was employed in this study to generate mean initial mathematics scores and average rates of mathematics achievement growth in Grades K–8. Nonetheless, it is quite possible that mathematics achievement growth might be significantly different for students with different levels of initial status (see Ai, 2002; Bodovski & Farkas, 2007). For example, students with minimal preschool academic experience may enter kindergarten with low mathematics skills, and respond well to early mathematics instruction thereby experiencing high rates of growth between kindergarten and first grade (i.e., low initial status with rapid growth). Alternatively, students who have low initial mathematics achievement scores may struggle to master early whole number concepts or receive ineffective instruction and demonstrate low rates of growth between kindergarten and first grade (i.e., low initial status with slow growth). Along these same lines, students with enriched preschool experiences may enter kindergarten with high early mathematics achievement and (a) experience slower growth through first grade (i.e., high initial status with slow growth) because there is less new material for these students to master, or (b) continue to demonstrate accelerated mathematics knowledge and maintain high rates of growth (i.e., high initial status with rapid growth). To this end, latent class growth analyses (Jung & Wickrama, 2008) could be conducted to evaluate the extent to which there are different patterns of mathematics achievement growth for different students based on their scores at kindergarten entry, and more precisely represent the variations in mathematics achievement growth.

Potential variations in the utility of self-efficacy as an important covariate of academic achievement also bear mentioning as an important limitation. Self-efficacy has been established as a reliable predictor of academic achievement, but there is a good deal of research that suggests that a vast array of intra and interpersonal skills predict academic achievement and postsecondary outcomes (Farrington et al., 2012). Whereas it was the aim of this study to explore the relationships between potential student survey indicators of mathematics self-efficacy and mathematics achievement scores, it may be more informative to consider the combined and complementary impact of a range of intra and interpersonal skills on mathematics achievement. For example, it could be that self-efficacy, conscientiousness, and persistence all work in concert such that conscientious students with high levels of mathematics self-efficacy demonstrate a propensity to persist in mathematics tasks and a higher likelihood of successfully completing mathematics assignments, which leads to improved rates of mathematics achievement growth in the middle grades. Similarly, it could be that grit or resourcefulness is particularly important intra and interpersonal skills that mediate the relationships between mathematics self-efficacy and mathematics achievement. Furthermore, these various intra and interpersonal skills may have variable relationships with mathematics self-efficacy and achievement for students from different demographic backgrounds. For example, low-SES students with high academic engagement, a strong internal locus of control, high self-esteem, and a positive attitude toward school have demonstrated more resilience in their mathematics learning as compared to their low-SES peers (Borman & Overman, 2004). In short, the current study examined only one small aspect of a complex system of intra and interpersonal skills and the ways in which these various personal skills and characteristics may interact with one another and ultimately operate as a cohesive system warrants consideration.

Lastly, the models examined here were able to explain a relatively small amount of the total variance in the self-efficacy measures in Grades 3, 5, and 8 (see Table 6). The limited number and quality of survey items for assessing mathematics self-efficacy may have contributed to these findings. Additional covariates should be examined to determine whether instructional factors and/or student interest or additional demographic characteristics could more fully explain the variance in mathematics self-efficacy. It could also be worthwhile to employ more targeted assessments and surveys of mathematics self-efficacy or administer additional items and conduct factor analyses to confirm the relevance of each item to the structure of the latent constructs being assessed.

Implications

Despite the limitations of the current study, the results of these analyses highlight a number of important areas of Grades K–8 mathematics efficacy and achievement growth with implications for both future research and intervention development. First, the rate at which young children acquire mathematics in kindergarten and first grade is much faster than the rate that they learn mathematics in the middle grades. The potential sources of this difference warrant exploration. Late elementary and middle school mathematics instruction, intervention practices, assessment validity, measure refinement, and progress monitoring are all plausible areas for investigation. Secondly, SES is an overwhelmingly influential factor in almost all aspects of mathematics learning and self-efficacy development. Students from above median SES households were nearly universally advantaged in both status and rates of growth for mathematics achievement and self-efficacy, particularly in the middle grades. Investigating home and school-level factors related to SES and providing targeted support and instruction to students from less affluent households may help to attenuate the effects of SES on achievement and begin to narrow this socioeconomic achievement gap. Future research on specific interventions is recommended.

Next, there are clear relationships between mathematics achievement growth and self-efficacy development; however, the mechanisms that link these two constructs remain unknown. Controlled research studies that track the effects of mathematics and intra and interpersonal skill interventions on both mathematics achievement and intra and interpersonal skill development can help identify these causal mechanisms. In turn, these studies may begin to illuminate effective intervention strategies to support students and successfully increase both mathematics competence and intra and interpersonal skills.

Lastly, there are sex differences in the system of mathematics achievement and self-efficacy. Although the differences detected here were quite small, they are substantively important based on continued evidence of differential postsecondary mathematics outcomes. If particular mathematics achievement markers are more highly associated with self-efficacy levels for females than for males, creating achievement goals around those pivotal points might be particularly effective for closing the sex gap in mathematics achievement and STEM-related outcomes. Furthermore, identifying the non-mathematics achievement factors that influence students who report high levels of self-efficacy might provide a starting point for intervention development.

In sum, the current study implies clear opportunities for better understanding and supporting growth in mathematics achievement and self-efficacy. Although the persistence of achievement differences based on demographic characteristics is disheartening, the findings discussed here suggest that mathematics instruction in Grades K–8 is ripe with potential for intervention. Based on the presence of early differences in achievement, early intervention programs that aim to level the playing field for young children by increasing achievement before kindergarten entry are extremely valuable. However, it should also be noted that the relationships between self-efficacy development and third grade achievement/growth in Grades 3–8 speaks to the importance of continued progress monitoring and targeted intervention throughout elementary and middle school. Ultimately, the results of this study suggest that instruction and interventions that facilitate improved mathematics achievement at any point in Grades K–8 can promote mathematics self-efficacy and better prepare students for the challenges of high school mathematics and beyond.

REFERENCES CITED

Achieve. (2008). *American diploma project algebra II end-of-course exam: 2008 annual report*. Washington, DC: Author.

Adelman, C. (2006). *The toolbox revisited: Paths to degree completion from high school through college.* Washington, DC: U.S. Department of Education.

Ai, X. (2002). Gender differences in growth in mathematics achievement: Three-level longitudinal and multilevel analyses of individual, home, and school influences. *Mathematical Thinking and Learning, 4*(1), 1-22.

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

Asparouhov, T. (2005). Sampling weights in latent variable modeling. *Structural Equation Modeling, 12*(3), 411-434.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory.* Englewood Cliffs, NJ: Prentice Hall.

Bandura, A. (1997). *Self-efficacy: The exercise of control.* New York: Freeman.

Beede, D., Julian, T., Langdon, D., McKittrick, G., Khan, B., & Doms, M. (2011). Women in STEM: A gender gap to innovation. *Economics and Statistics Administration Issue Brief, 4,* 1-11.

Berends, M., Lucas, S. R., Sullivan, T., & Briggs, R. J. (2005). *Examining gaps in mathematics achievement among racial-ethnic groups, 1972-1992*. Santa Monica, CA: Rand Corporation.

Bloom, H. S., Hill, C. J., Black, A. R., & Lipsey, M. W. (2008). Performance trajectories and performance gaps as achievement effect-size benchmarks for educational interventions. *Journal of Research on Educational Effectiveness, 1*(4), 289-328.

Blum, R. W., & Libbey, H. P. (2004). School connectedness—Strengthening health and education outcomes for teenagers. *Journal of School Health, 74,* 229–299.

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

Burchinal, M., McCartney, K., Steinberg, L., Crosnoe, R., Friedman, S. L., McLoyd, V., & Pianta, R. (2011). Examining the black–white achievement gap among low income children using the NICHD study of early child care and youth development. *Child Development, 82*(5), 1404-1420.

Burkam, D. T., Ready, D. D., Lee, V. E., & LoGerfo, L. F. (2004). Social-class differences in summer learning between kindergarten and first grade: Model specification and estimation. *Sociology of Education, 77*(1), 1-31.

Byrne, B. M. (2012). *Structural equation modeling with Mplus: Basic concepts, applications, and programming.* New York, NY: Routledge.

Camara, W. J. (2005). Broadening predictors of college success. In W. Camara & E. W. Kimmel (Eds.), *Choosing students: Higher education admissions tools for the 21st century*, (pp. 81-105), New York, NY: Routledge.

Caprara, G. V., Vecchione, M., Alessandri, G., Gerbino, M., & Barbaranelli, C. (2011). The contribution of personality traits and self-efficacy beliefs to academic achievement: A longitudinal study. *British Journal of Educational Psychology*, *81*(1), 78-96.

Claessens, A., Duncan, G., & Engel, M. (2009). Kindergarten skills and fifth-grade achievement: Evidence from the ECLS-K. *Economics of Education Review*, *28*(4), 415-427.

Claessens, A., & Engel, M. (2013). How important is where you start? Early mathematics knowledge and later school success. *Teachers College Record, 115,* 1-29.

Common Core State Standards Initiative. (2010). *Common core state standards for mathematics*. Retrieved from Common Core State Standards Initiative website: http://www.corestandards.org/.

Cooper, H., Nye, B., Charlton, K., Lindsay, J., & Greathouse, S. (1996). The effects of summer vacation on achievement test scores: A narrative and meta-analytic review. *Review of Educational Research, 66*(3), 227-268.

Crane, J. (1996). Effects of home environment, SES, and maternal test scores on mathematics achievement. *Journal of Educational Research, 89*(5), 305-314.

Cutuli, J. J., Desjardins, C. D., Herbers, J. E., Long, J. D., Heistad, D., Chan, C. K., ... & Masten, A. S. (2013). Academic achievement trajectories of homeless and highly mobile students: Resilience in the context of chronic and acute risk. *Child Development, 84*(3), 841-857.

Devine, A., Fawcett, K., Szucs, D., & Dowker, A. (2012). Gender differences in mathematics anxiety and the relation to mathematics performance while controlling for test anxiety. *Behavioral and Brain Functions, 8*(33), 2-9.

Di Giunta, L., Alessandri, G., Gerbino, M., Luengo Kanacri, P., Zuffiano, A., & Caprara, G. V. (2013). The determinants of scholastic achievement: The contribution of personality traits, self-esteem, and academic self-efficacy. *Learning and Individual Differences, 27*, 102-108.

Ding, C. S., & Davison, M. L. (2005). A longitudinal study of math achievement gains for initially low achieving students. *Contemporary Educational Psychology, 30*(1), 81-95.

Diseth, Å. (2011). Self-efficacy, goal orientations and learning strategies as mediators between preceding and subsequent academic achievement. *Learning and Individual Differences*, *21*(2), 191-195.

Downey, D.B., von Hippel, P.T., Hughes, M. (2008). Are "failing" schools really failing? Using seasonal comparisons to evaluate school effectiveness. *Sociology of Education, 81*(3), 242–270.

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*(6), 1428–1446.

Duncan, G. J., & Magnuson, K. A. (2005). Can family socioeconomic resources account for racial and ethnic test score gaps? *The Future of Children*, *15*(1), 35-54.

Duncan, G. J., & Murnane, R. J. (Eds.). (2011). *Whither opportunity? Rising inequality, schools, and children's life chances*. New York: Russell Sage Foundation.

Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D. & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development, 82*(1), 405-432.

Dweck, C. S. (2008). *Mindsets and math/science achievement.* New York, NY: Carnegie Corp. of New York–Institute for Advanced Study Commission on Mathematics and Science Education.

Dweck, C.S., Walton, G.M., and Cohen, G.L. (2011). *Academic tenacity: Mindsets and skills that promote long-term learning.* White paper prepared for the Gates Foundation. Seattle, WA.

Eccles, J. S., & Midgley, C. (1989). Stage-environment fit: Developmentally appropriate classrooms for young adolescents. In C. Ames & R. Ames (Eds.), *Research on motivation in education* (Vol. 3, pp. 139-186). San Diego: Academic Press.

Eisenberg, N., Martin, C. L., & Fabes, R. A. (1996). Gender development and gender effects. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 358-396). New York: Macmillan.

Farrington, C. A., Roderick, M., Allensworth, E., Nagaoka, J., Keyes, T. S., Johnson, D. W., & Beechum, N. O. (2012). *Teaching adolescents to become learners: The role of noncognitive factors in shaping school performance--A critical literature review*. Chicago, IL: Consortium on Chicago School Research.

Fennema, E., Peterson, P. L., Carpenter, T. P., & Lubinski, C. A. (1990). Teachers' attributions and beliefs about girls, boys, and mathematics. *Educational Studies in Mathematics, 21*(1), 55-69.

Foegen, A. (2008). Progress monitoring in middle school mathematics: Options and issues. *Remedial and Special Education, 29*(4), 195-207.

Fortin, N. M. (2008). The gender wage gap among young adults in the United States: The importance of money vs. people. *The Journal of Human Resources, XLIII*(4), 884-918.

Fredricks, J. A., & Eccles, J. S. (2002). Children's competence and value beliefs from childhood through adolescence: growth trajectories in two male-sex-typed domains. *Developmental Psychology, 38*(4), 519.

Frenzel, A. C., Pekrun, R., & Goetz, T. (2007). Girls and mathematics—A “hopeless” issue? A control-value approach to gender differences in emotions towards mathematics. *European Journal of Psychology of Education, 22*(4), 497-514.

Gamoran, A. & Hannigan, E. C. (2000). Algebra for everyone? Benefits of college-preparatory mathematics for students with diverse abilities in early secondary school. *Educational Evaluation and Policy Analysis*, *22*(3), 241-254.

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

Ginsburg, H., & Baroody, A. J. (1983). *The test of early mathematics ability.* Austin, Texas: Pro-ed.

Graham, S. (1994). Motivation in African Americans. *Review of Educational Research, 64*, 55-117.

Gunderson, E. A., Ramirez, G., Levine, S. C., & Beilock, S. L. (2012). The role of parents and teachers in the development of gender-related math attitudes. *Sex Roles, 66*(3-4), 153-166.

Hair, E., Halle, T., Terry-Humen, E., Lavelle, B., & Calkins, J. (2006). Children's school readiness in the ECLS-K: Predictions to academic, health, and social outcomes in first grade. *Early Childhood Research Quarterly*, *21*(4), 431-454.

Heckman, J. J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science, 312*(5782), 1900-1902.

Heckman, J. J., & Rubinstein, Y. (2001). The importance of noncognitive skills: Lessons from the GED testing program. *American Economic Review, 91*, 145-149.

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational and Behavioral Statistics, 6*(2), 107-128.

Howard, T. C. (2010). *Why race and culture matter in schools: Closing the achievement gap in America's classrooms.* New York: Teachers College Press.

Hyde, J. S., Fennema, E., Ryan, M., Frost, L. A., & Hopp, C. (1990). Gender comparisons of mathematics attitudes and affect. *Psychology of Women Quarterly, 14*(3), 299-324.

IBM Corp. (2011). *IBM SPSS Statistics for MacIntosh, Version 20.0.* Armonk, NY: IBM Corp.

Jacob, B. A. (2002). Where the boys aren't: Non-cognitive skills, returns to school and the gender gap in higher education. *Economics of Education Review*, *21*(6), 589-598.

Jacobs, J. E., Lanza, S., Osgood, D. W., Eccles, J. S., & Wigfield, A. (2002). Changes in children’s self-competence and values: Gender and domain differences across grades one through twelve. *Child Development, 73*(2), 509-527.

Jencks, C., & Phillips, M. (Eds.). (2011). *The black-white test score gap.* Washington, DC: Brookings Institution Press.

Joët, G., Usher, E. L., & Bressoux, P. (2011). Sources of self-efficacy: An investigation of elementary school students in France. *Journal of Educational Psychology, 103*, 649-663.

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

Judge, T. A., & Hurst, C. (2008). How the rich (and happy) get richer (and happier): relationship of core self-evaluations to trajectories in attaining work success. *Journal of Applied Psychology, 93*(4), 849-863.

Jung, T., & Wickrama, K. A. S. (2008). An introduction to latent class growth analysis and growth mixture modeling. *Social and Personality Psychology Compass, 2*(1), 302-317.

Kao, G., & Thompson, J. S. (2003). Racial and ethnic stratification in educational achievement and attainment. *Annual Review of Sociology*, *29*, 417-442.

Keller, C. (2001). Effect of teachers' stereotyping on students' stereotyping of mathematics as a male domain. *The Journal of Social Psychology, 141*(2), 165-173.

Keller-Margulis, M. A., Mercer, S. H., & Shapiro, E. S. (2012). Differences in growth on math curriculum-based measures using triannual benchmarks. *Assessment for Effective Intervention, 39*(3), 146-155.

Kim, J. (2004). Summer reading and the ethnic achievement gap. *Journal of Education for Students Placed at Risk, 9*(2), 169-188.

Kitsantas, A., Cheema, J., & Ware, H. W. (2011). Mathematics achievement: The role of homework and self-efficacy beliefs*.* *Journal of Advanced Academics, 22*(2), 310-339.

Kline, R. B. (2010). *Principles and practice of structural equation modeling (3rd Ed.).* New York: Guilford Press.

Larson, L. M., Pesch, K. M., Surapaneni, S., Bonitz, V. S., Wu, T. F., & Werbel, J. D. (2014). Predicting graduation: The role of mathematics/science self-efficacy. *Journal of Career Assessment, 1,* 1-11.

Lee, J. (2004). Multiple facets of inequity in racial and ethnic achievement gaps. *Peabody Journal of Education, 79*(2), 51-73.

Lee, J. (2012). College for all gaps between desirable and actual P–12 math achievement trajectories for college readiness. *Educational Researcher, 41*(2), 43-55.

Lee, J., & Stankov, L. (2013). Higher-order structure of noncognitive constructs and prediction of PISA 2003 mathematics achievement. *Learning and Individual Differences, 26,* 119-130.

Lee, V. E. & Burkam, D. T. (2002). *Inequality at the starting gate: Social background differences in achievement as children begin school.* Washington, DC: Economic Policy Institute.

Li, Q. (1999). Teachers' beliefs and gender differences in mathematics: A review. *Educational Research, 41*(1), 63-76.

Linver, M. R. & Davis-Kean, P. E. (2005). The slippery slope: What predicts math grades in middle and high school? *New Directions for Child and Adolescent Development,110*, 49-64.

Lippman, L., Burns, S., & McArthur, E. (1996). *Urban schools: The challenge of location and poverty.* Washington, DC: U.S. Department of Education, National Center for Education Statistics.

Long, M. C., Iatarola, P., & Conger, D. (2009). Explaining gaps in readiness for college-level math: The role of high school courses. *Education, 4*(1), 1-33.

Lubienski, S. T. (2002). A closer look at Black-White mathematics gaps: Intersections of race and SES in NAEP achievement and instructional practices data. *Journal of Negro Education, 71*(4), 269-287.

Marsh, H. W. (1992). *Self Description Questionnaire (SDQ) II: A theoretical and empirical basis for the measurement of multiple dimensions of adolescent self-concept. A test manual and research monograph.* New South Wales, Australia: University of Western Sydney.

Martinot, D., Bagès, C., & Désert, M. (2012). French children's awareness of gender stereotypes about mathematics and reading: When girls improve their reputation in math. *Sex Roles, 66*(3-4), 210-219.

Mattern, K. D., & Shaw, E. J. (2010). A look beyond cognitive predictors of academic success: Understanding the relationship between academic self-beliefs and outcomes. *Journal of College Student Development, 51*(6), 665-678.

Meltzer, D. E. (2002). The relationship between mathematics preparation and conceptual learning gains in physics: A possible "hidden variable" in diagnostic pretest scores. *American Journal of Physics*, *70*, 1259–1268.

Midgley, C., Feldlaufer, H., & Eccles, J. S. (1989). Student/teacher relations and attitudes toward mathematics before and after the transition to junior high school. *Child Development, 60*(4), 981-992.

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

Morgan, S. L., Gelbgiser, D., & Weeden, K. A. (2013). Feeding the pipeline: Gender, occupational plans, and college major selection. *Social Science Research, 42*, 989-1005.

Murnane, R. J., Willett, J. B., Bub, K. L., & McCartney, K. (2006). Understanding trends in the black-white achievement gaps during the first years of school. *Brookings-Wharton Papers on Urban Affairs.*

Muthén, L. K., & Muthén, B. O. (2013). *Mplus User's Guide, v 7.1.* Los Angeles, CA.: Muthén and Muthén, UCLA.

Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Math= male, me= female, therefore math≠ me. *Journal of Personality and Social Psychology, 83*(1), 44-59.

Nosek, B. A., & Smyth, F. L. (2011). Implicit social cognitions predict sex differences in math engagement and achievement. *American Educational Research Journal, 48*(5), 1125-1156.

Nowell, A., & Hedges, L. V. (1998). Trends in gender differences in academic achievement from 1960 to 1994: An analysis of differences in mean, variance, and extreme scores. *Sex Roles, 39*(1-2), 21-43.

Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research, 66*(4), 543-578.

Pajares, F., & Miller, M. D. (1994). Role of self-efficacy and self-concept beliefs in mathematical problem solving: A path analysis. *Journal of Educational Psychology, 86*(2), 193-203.

Pajares, F., Miller, M. D., & Johnson, M. J. (1999). Gender differences in writing self-beliefs of elementary school students. *Journal of Educational Psychology, 91*, 50-61.

Pajares, F., & Schunk, D. (2001). The development of academic self-efficacy. In A. Wigfield & J. Eccles (Eds.) *Development of achievement motivation.* San Diego: Academic Press.

Pellegrino, J. W. & Hilton, M. L. (Eds.). (2012). *Education for life and work: Developing transferable knowledge and skills in the 21st* century. Washington, DC: National Academies Press.

Perez-Felkner, L., McDonald, S. K., Schneider, B., & Grogan, E. (2012). Female and male adolescents' subjective orientations to mathematics and the influence of those orientations on postsecondary majors. *Developmental Psychology, 48*(6), 1658-1673.

Pintrich, P. R., & Schunk, D. H. (1996). *Motivation in education: Theory, research, and applications.* Englewood Cliffs, NJ: Merrill.

Reardon, S. F. (2011). The widening academic achievement gap between the rich and the poor: New evidence and possible explanations. In Greg J. Duncan & Richard J. Murnane (Eds.). *Whither opportunity? Rising inequality, schools, and children's life chances,* (pp. 91-116). New York: Russell Sage Foundation.

Reardon, S. F., & Galindo, C. (2009). The Hispanic-White achievement gap in math and reading in the elementary grades. *American Educational Research Journal*, *46*(3), 853-891.

Reardon, S. F., Robinson-Cimpian, J. P., & Weathers, E. S. (2014). Patterns and trends in racial/ethnic and socioeconomic academic achievement gaps. In Helen A. Ladd & Margaret E. Goertz (Eds.). *Handbook of research in education finance and policy*. Hillsdale, NJ: Lawrence Erlbaum.

Reckase, M. D. (2004). The real world is more complicated than we would like. *Journal of Educational and Behavioral Statistics, 29*(1) 117-120.

Robinson, J. P., & Lubienski, S. T. (2011). The development of gender achievement gaps in mathematics and reading during elementary and middle school: Examining direct cognitive assessments and teacher ratings. *American Educational Research Journal, 48*(2), 268-302.

Rock, D. A. & Pollack, J. M. (2002). *Early Childhood Longitudinal Study-Kindergarten Class of 1998–99 (ECLS–K), Psychometric Report for Kindergarten Through First Grade, NCES 2002–05*. Washington, DC: NCES.

Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika, 66*(4), 507-514.

Schmader, T., Johns, M., & Barquissau, M. (2004). The costs of accepting gender differences: The role of stereotype endorsement in women's experience in the math domain. *Sex Roles, 50*(11-12), 835-850.

Schunk, D. H., & Meece, J. L. (2006). Self-efficacy development in adolescence. In Frank Pajares & Tim Urdan (Eds.). *Self-Efficacy beliefs of adolescents,* (pp. 71-96). Scottsdale, AZ: Information Age.

Shams, F., Mooghali, A. R., & Soleimanpour, N. (2011). The mediating role of academic self-efficacy in the relationship between personality traits and mathematics performance. *Procedia-Social and Behavioral Sciences, 29*, 1689-1692.

Siegler, R. S., Duncan, G. J., Davis-Kean, P. E., Duckworth, K., Claessens, A., Engel, M., ... & Chen, M. (2012). Early predictors of high school mathematics achievement. *Psychological Science, 23*(7), 691-697.

Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research, 75*(3), 417-453.

Smith, E. P., Walker, K., Fields, L., Brookins, C. C., & Seay, R. C. (1999). Ethnic identity and its relationship to self-esteem, perceived efficacy and prosocial attitudes in early adolescence. *Journal of Adolescence, 22*(6), 867-880.

Stapleton, L. M. (2008). Variance estimation using replication methods in structural equation modeling with complex sample data. *Structural Equation Modeling, 15*(2), 183-210.

Tourangeau, K., Nord, C., Le, T., Sorongon, A. G., & Najarian, M. (2009). *Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K): Combined user's manual for the ECLS-K eighth-grade and K-8 full sample data files and electronic codebooks* (NCES 2009-004). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education, Washington, DC.

Tracey, T. & Robbins, S. (2004). *Ethnic and gender differences in interests and academic skills: A longitudinal examination of grades 8 through 12.* Iowa City, IA: ACT.

Trusty, J., Robinson, C. R., Plata, M., & Ng, K. M. (2000). Effects of gender, socioeconomic status, and early academic performance on postsecondary educational choice. *Journal of Counseling & Development, 78*(4), 463-472.

U.S. Department of Education, (2010). *A blueprint for reform: The reauthorization of the elementary and secondary education act.* Retrieved from: www2.ed.gov/policy/elsec/leg/blueprint

Vanneman, A., Hamilton, L., Anderson, J. B., & Rahman, T. (2009). *Achievement gaps: How black and white students in public schools perform in mathematics and reading on the National Assessment of Educational Progress.* Statistical Analysis Report. NCES 2009-455. National Center for Education Statistics.

Whaley, A. L., & Noel, L. (2011). Sociocultural theories, academic achievement, and African American adolescents in a multicultural context: a review of the cultural incompatibility perspective. *Social Psychology of Education, 14*(2), 149-168.

White, K. R. (1982). The relation between socioeconomic status and academic achievement. *Psychological Bulletin, 91*(3), 461.

Wigfield, A., Eccles, J. S., & Pintrich, P. R. (1996). Development between the ages of 11 and 25. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 148-185). New York: Macmillan.

Wimberly, G. L. & Noeth, R. J. (2005). *College readiness begins in middle school.* Washington, DC: ACT.