Title

The Role of Mediators in the Development of Longitudinal Mathematics Achievement Associations

Journal
Child Development, 86(6)

ISSN
0009-3920

Authors
Watts, TW
Duncan, GJ
Chen, M
et al.

Publication Date
2015-11-01

DOI
10.1111/cdev.12416

Peer reviewed

CHILD DEVELOPMENT

Child Development, xxxx 2015, Volume 00, Number 0, Pages 1–16

The Role of Mediators in the Development of Longitudinal Mathematics Achievement Associations

Tyler W. Watts and Greg J. Duncan
University of California, Irvine

Meichu Chen
University of Michigan

Amy Claessens
University of Chicago

Pamela E. Davis-Kean
University of Michigan

Kathryn Duckworth
University of London

Mimi Engel
Vanderbilt University

Robert Siegler
Carnegie Mellon University and Siegler Center for Innovative Learning, Beijing Normal University

Maria I. Susperreguy
Pontificia Universidad Católica de Chile

Despite research demonstrating a strong association between early and later mathematics achievement, few studies have investigated mediators of this association. Using longitudinal data (n = 1,362), this study tested the extent to which mathematics self-concepts, school placement, executive functioning, and proficiency in fractions and division account for the association between mathematics achievement in first grade and at age 15. As hypothesized, a strong longitudinal association between first-grade and adolescent mathematics achievement was present (β = .36) even after controlling for a host of background characteristics, including cognitive skills and reading ability. The mediators accounted for 39% of this association, with mathematics self-concept, gifted and talented placement, and knowledge of fractions and division serving as significant mediators.

Duncan et al. (2007) analyzed six longitudinal studies in order to assess the predictive power of a number of school-entry measures of achievement and socioemotional skills. They found that after controlling for a variety of prior child and family background characteristics, mathematics knowledge at school entry appeared to be the strongest predictor of later mathematics achievement. This finding was robust across the six data sets. However, important questions regarding the underlying mechanisms that produced these predictive relations remained. These questions motivated the current article's mediational approach to modeling the association between early and later mathematics achievement. By pursuing a comprehensive investigation of a set of mediators, we may better understand the pathways that contribute to stability over time of individual differences in mathematics knowledge.

Models of Mathematics Achievement

No current theoretical model incorporates the wide array of factors likely to contribute to long-run mathematics achievement. Consequently, most empirical investigations of mathematics development focus on one or two specific factors, such as

The authors are grateful to the NSF-supported Center for the Analysis of Pathways from Childhood to Adulthood (Grant 0322356), the NICHD-supported Irvine Network on Interventions in Development (Grant HD065704 P01), and the Institute of Educational Sciences (Grant R324C100004) for research support. The authors thank Ana Auger, Drew Bailey, Doug Clements, George Farkas, David Geary, Katerina Schenke, Deborah Vandell, and the NICHD Early Child Care Research Network for their helpful contributions to this project. The authors also thank the participants of the NICHD SECCYD.

Correspondence concerning this article should be addressed to Tyler W. Watts, School of Education, University of California, Irvine, 2031 Education Building, Irvine, CA 92697-5500. Electronic mail may be sent to twatts@uci.edu.

0009-3920/2015/xxxx-xxxx
DOI: 10.1111/cdev.12416

motivation (e.g., Marsh, Byrne, & Yeung, 1999), executive function (e.g., Blair & Razza, 2007), or math-specific skills (e.g., Siegler et al., 2012). Although such examinations provide information about specific correlates of mathematics achievement, more encompassing analyses are needed if we are to understand how a theoretically diverse set of math-related factors act in concert. After all, as children develop mathematical skills, a set of cognitive, socioemotional, motivational, and environmental characteristics all operate simultaneously. Thus, a comprehensive analytic approach needs to incorporate a broad range of potentially important variables.

Our selection of possible mediators was guided by a broad bioecological perspective (Bronfenbrenner & Morris, 2006) that led us to model mathematics achievement as a function of both personal and environmental characteristics. Furthermore, in order to design our empirical model, we set forth with a criterion that guided our selection of each mediator. First, candidate mediators needed to have a straightforward theoretical link to mathematics achievement. Second, this theoretical link needed to be supported by convincing empirical evidence. Third, the potential mediators must have been measured quantitatively in a manner that allowed for investigation in a longitudinal analysis. Fortunately, we were able to use a data set that provided timely measures of all but one of our desired set of mediators.

These criteria led us to simultaneously consider four distinct domains, each of which has been shown individually to influence mathematics achievement, but which have not been considered together. On the basis of bioecological framework, we first examined characteristics and skills of the child and then moved outward to the environmental context. We began with the child's mathematical knowledge. Because mathematics is a hierarchical area in which later knowledge incorporates and builds on earlier knowledge, we hypothesized that certain mathematics skills thought to act as gatekeepers to higher level mathematics achievement would be important mediators between first-grade and age 15 mathematical skills. The second mediator consists of the domain-general cognitive skills covered by the executive functioning (EF) umbrella, as these intellectual processes have been found to predict mathematics achievement, above and beyond earlier mathematical knowledge (Fuhs, Nesbitt, Farran, & Dong, 2014).

The third potential mediator that we examined involved motivation. Our main index of motivation was self-concept regarding mathematics ability (on the logic that higher self-concepts regarding math ability both reflect and inspire greater motivation to exercise this ability). Math self-concept has been shown to be related to the development of mathematics achievement (e.g., Bong & Skaalvik, 2003). In particular, children who perceive themselves as good at mathematics typically fare better on subsequent examinations.

Finally, moving from inside to outside the child, the fourth potential mediator that we examined was school and classroom environments. In particular, school placements into gifted and talented programs and special education have been linked to subsequent mathematics achievement (Bhatt, 2009; Morgan, Frisco, Farkas, & Hibel, 2010), and operate by taking children out of their regular classrooms to provide specialized instruction in mathematics.

We hypothesized that each of these four factors could substantially mediate the association between early and later mathematics achievement. However, we recognize that these four potential mediators are not the only factors that influence the growth of mathematical knowledge. Quality of classroom instruction and teacher characteristics were among the variables we hoped to pursue, but empirical limitations did not allow us to investigate them. Although the present study is not fully comprehensive, it does include a diverse set of mediators that encompass math-specific competencies, domain-general cognitive skills, motivation, and school-related structures. We know of no other study that has examined these four factors in one analysis, even though this comprehensive approach more closely resembles the complex nature of how these variables influence development.

Our empirical work extends previous longitudinal analyses (e.g., Claessens & Engel, 2013; Duncan et al., 2007; Watts, Duncan, Siegler, & Davis-Kean, 2014) by modeling age 15 mathematics achievement as a function of first-grade academic and cognitive skills, using data drawn from the National Institute of Child Health and Human Development (NICHD) Study of Early Childcare and Youth Development (SECCYD). We control for a host of important variables, all measured at or before first grade, including child temperament, socioemotional skills, cognitive functioning, and family and child background characteristics. We then assess the extent to which the associations between early and later mathematics achievement can be accounted for by our mediational pathways. In the following sections, we briefly review literature on each mediator, and discuss implications of the study for theory and practice.

Background

Fractions and Division Knowledge

Because of the hierarchical structure of mathematical skill development, prior knowledge of specific mathematical competencies is crucial for the development of later mathematics achievement. Although a variety of mathematical skills have been implicated in the development of long-run mathematics achievement (e.g., math reasoning, Nunes, Bryant, Barros, & Sylva, 2012; number system knowledge, Geary, Hoard, Nugent, & Bailey, 2013; numeracy skills, Jordan, Kaplan, Ramineni, & Locuniak, 2009), knowledge of fractions and division appear to be particularly crucial. Siegler and colleagues hypothesized that understanding fractions plays a central role in developing high school mathematics capacity (Siegler, Thompson, & Schneider, 2011; Siegler et al., 2012). If first-grade mathematics achievement is strongly associated with the development of fraction and division skills, and if these skills are a gateway to algebraic proficiency, then fractions and division knowledge should mediate the relation between first-grade mathematics achievement and adolescent mathematical ability.

When students encounter fractions, they are forced to reorganize their understanding of the properties of numbers, from one that applies only to whole numbers to a more general understanding. Unlike whole numbers, fractions do not consistently increase when multiplied, do not consistently decrease when divided, do not have unique successors, and do not have unique symbolic representations. Gaining a broader understanding of what all real numbers have in common (e.g., that they can represent magnitudes) is crucial for success in high school mathematics (Martin, Strutchens, & Elliott, 2007). Solving even simple algebraic equations (e.g., $2/3X = 8$) is a formidable task if students do not understand the basic properties of fractions (e.g., that X will be larger than 8).

The examination of two longitudinal data sets by Siegler et al. (2012) revealed that fractions and division knowledge in fifth grade were uniquely predictive of high school algebra proficiency, even after accounting for domain-general cognitive functioning, reading ability, and whole number skills. Although both division and fraction skills were found to be predictive, they are closely related concepts, as a fraction is a representation of a division problem (e.g., 3/4 can be thought of as 3 divided by 4). Furthermore, fifth graders' whole number addition, subtraction, and multiplication were less predictive of high school algebraic ability. Similar shorter term relations between fraction knowledge and later mathematics learning have been reported elsewhere (Booth & Newton, 2012), and Bailey, Siegler, and Geary (2014) found that first-grade proficiency with whole numbers predicted eighth-grade proficiency with fractions.

Executive Function

Another mediator that could account for the longitudinal relation between early and later mathematics achievement is EF—higher order cognitive and regulatory processes that involve working memory, inhibitory control, attention regulation, set shifting, and complex planning (e.g., Clark, Pritchard, & Woodward, 2010; Fuhs et al., 2014). If early mathematical proficiency leads students to practice EF skills, then we may detect an association between early mathematics achievement and growth in EF. Furthermore, if these EF skills play a role in developing later mathematical skills, then EF should account for some portion of the association between early and later mathematics achievement.

EF skills develop during early childhood and grow over the course of childhood and adolescence (Gathercole, Pickering, Ambridge, & Wearing, 2004). It is not difficult to imagine how EF could play a key mediational role in the long-term development of mathematical skills. For example, consider the cognitive processes required to solve a fractions addition problem that calls for derivation of the least common denominator. The student would need high inhibitory control to resist the temptation to treat the fractions as whole numbers and add numerators and denominators separately. The student would also need working memory and attention regulation skills to hold potential solutions in mind, while considering possible common denominators. Set shifting would be needed to alternate focus between the two denominators, and cognitive flexibility would be important in finding the lowest multiple common of the denominators. Finally, as students become more experienced solving such problems, they may develop more advanced strategies that require a certain amount of forethought and planning (e.g., prime factorization).

Indeed, research investigating concurrent associations between EF and mathematical skills have found fairly strong correlations (e.g., Blair & Razza, 2007; St. Clair-Thompson & Gathercole, 2006), but far less research has investigated the predictive relation between EF and subsequent measures of mathematics achievement. Recent studies suggest that such an association may exist. Clark et al. (2010) found that a host of EF skills measured during preschool, including set shifting and inhibitory control, positively predicted mathematics achievement measured 2 years later. This result held after controlling for general cognitive ability and reading achievement.

Studies have also found longitudinal associations between early measures of working memory and math achievement measured in childhood (Mazzocco & Kover, 2007; Welsh, Nix, Blair, Bierman, & Nelson, 2010) and adolescence (Watts et al., 2014). Similar findings have been reported when EF skills have been examined as one latent construct. Most recently, Fuhs et al. (2014) used a single factor design to measure a broad array of EF skills-during preschool, including inhibitory control, working memory, and attention flexibility, finding that these competencies predicted mathematical skill growth during kindergarten.

For EF to qualify as a mediator of the association between early and later mathematics achievement, early mathematics ability must also lead to growth in subsequent EF. This might occur if those who excel academically from an early age seek new situations that further stimulate their cognitive growth (Fuhs et al., 2014). Consistent with this hypothesis, Welsh et al. (2010) found that academic skills at preschool entry were correlated with end-of-preschool working memory and attention control. However, Fuhs et al. (2014) found that while mathematical skills measured at the beginning of preschool predicted end-of-preschool EF, this relation did not hold when EF was measured at the end of kindergarten. These conflicting findings indicate a need for longer term longitudinal studies of the relation between early mathematical skills and later EF.

Self-Concept

Self-concept of ability (SCA), which describes an individual's knowledge and perceptions regarding his or her academic capabilities (Bong & Skaalvik, 2003), has been hypothesized to influence students' motivation and affect toward engaging in particular subjects (Marsh et al., 1999). If a student with high mathematics achievement favorably compares herself with lower achieving peers, she may begin to view herself as especially competent in mathematics. This perception could also arise through positive feedback from parents and teachers. In turn, this enhanced self-concept could lead her to seek situations in which her self-perception would be affirmed, such as through taking more advanced mathematics classes (Bong & Skaalvik, 2003). Thus, self-perceptions that reflect responses to early-grade achievement patterns in mathematics could account for the relation between early and later mathematics achievement if students internalize these perceptions, which in turn alter students' approach to mathematics in a way that either enhances or diminishes their acquisition of new mathematical skills (Marsh & Craven, 2006).

Findings from empirical research investigating the link between SCA and mathematics achievement have produced mixed results. It has been hypothesized that self-perceptions arise in domain-specific contexts (e.g., mathematics), as students evaluate their performance in specific subjects against their peers (Bong & Skaalvik, 2003; Marsh & Craven, 2006). Over time, self-perceptions across domains may even act in opposition to one another. Indeed, Marsh, Byrne, and Shavelson (1988) found that verbal and mathematical self-concepts of high school students were uncorrelated, and although mathematics achievement positively predicted subsequent mathematics SCA, it negatively predicted verbal SCA. Moreover, although the domain-specific link between mathematics achievement at Time 1 and SCA at Time 2 has been observed in both high school (Marsh & Yeung, 1997; Marsh et al., 1999) and primary school (Skaalvik & Valås, 1999) students, it remains unclear whether SCA continues to influence subsequent mathematics achievement. Multiple studies have failed to detect such an association in mathematics (Helmke & van Aken, 1995; Skaalvik & Valås, 1999), though other longitudinal investigations have found SCA to be a significant pathway to math achievement (Marsh & Yeung, 1997). Some have argued that the failures to detect a predictive link could be due to methodological shortcomings (Marsh et al., 1999), indicating a need for more rigorous empirical work investigating the longitudinal associations between self-concept and mathematics achievement.

School Placement

Another possible mediator of relations between first-grade and adolescent achievement is placement into special education or gifted and talented programs. This pathway would be supported if early achievement gains alter the chance of placement into gifted and talented programs or special education services, and if these school structures in turn boost or hinder subsequent mathematics achievement gains. Such relations could arise through several different mechanisms.

First, placement into special school programs could expose students to either higher or lower quality instruction. Some research has suggested that students placed into gifted and talented programs received higher quality instruction and better curricula (Bui, Craig, & Imberman, 2011), though the effect of such placement on achievement was small. Similarly, some investigators have argued that students placed into special education programs may receive more empirically supported and individualized instruction (Fuchs & Fuchs, 1998), but it is unknown if these techniques are consistently implemented across schools (Cook & Schirmer, 2003). Furthermore, special education programs have been criticized for subjecting students to lower expectations (Horn & Tynan, 2001), and recently there has been considerable pressure to move students with learning disabilities into "mainstream education" (Cook & Schirmer, 2003). School placements could also alter achievement trajectories through peer effects. To the extent that academically stronger peers improve the achievement of classmates (e.g., Imberman, Kugler, & Sacerdote, 2012), time spent in gifted and talented programs should support learning, and time spent in special education should impair it.

Some longitudinal evidence supports our hypothesis that placement into special academic programs mediates the relation between early and later mathematics achievement. Early grade mathematics achievement predicts selection into both gifted and talented programs (McClain & Pfeiffer, 2012) and special education services (Hibel, Farkas, & Morgan, 2010). However, the effect of these programs on subsequent mathematics achievement has not been established.

Special education programs typically aim to provide students specialized services to help mitigate the impact of their disability (Morgan et al., 2010). Some studies indicate that students with disabilities benefit from such services (Blackorby & Wagner, 1996; Hanushek, Kain, & Rivkin, 2002), but there is mounting evidence to the contrary. Morgan and colleagues' (2010) analysis of nationally representative data from the Early Childhood Longitudinal Study's Kindergarten Cohort found that placement into special education services had a negative association with subsequent mathematics achievement in some, but not all, model specifications. Similarly, Lane, Wehby, Little, and Cooley (2005) found that children assigned to special education classes often had lower scores in both mathematics and reading at the end of the school year than they did before receiving special education services.

Studies investigating the effects of placement into gifted and talented programs have also produced mixed results. Vaughn, Feldhusen, and Asher's (1991) meta-analysis of correlational studies examining the effect of "pull-out" programs, which periodically remove academically gifted students from their regular classes to receive enrichment, reported that such programs had a positive effect size of 0.65 *SD* (*SE* = 0.19) on student achievement. However, little research has taken selection bias into account, and strong experimental evidence remains virtually nonexistent.

Quasi-experimental studies have yielded conflicting results. Bui et al. (2011) employed a regression discontinuity approach to examine the effect of gifted and talented placement on middle school students; they found no impact on mathematics achievement. In contrast, Bhatt (2009) used instrumental variables to approximate the causal impact of gifted and talented placement on math achievement among middle and high school students, and found a large effect of program placement on mathematics achievement. Thus, substantial questions remain regarding the impact of gifted and talented program placement on student achievement.

Current Study

The current study investigates the extent to which mathematics SCA, EF, school placements, and fraction and division knowledge mediate the association between first-grade and age 15 mathematics achievement. It adds to previous research in several ways. First, we draw on multisite national data that include well-validated measures of achievement, allowing us to study achievement trajectories in unusual detail from first grade through the beginning of high school. Second, we employ an extensive set of control variables, including measures of socioemotional functioning and problem behaviors, cognitive functioning, and individual and familial background characteristics. Inclusion of these control variables should help to reduce bias in the estimation of the unique contributions that each proposed mediator makes to the longitudinal mathematics achievement trajectories we investigate.

Finally, we simultaneously test multiple hypotheses regarding underlying mechanisms that could account for the robust association reported between early and later math skills. Instead of taking the more common piecemeal approach that considers each mediator individually, our analysis estimates the relative contribution of each of the mediators controlling for the others—a strategy that should reduce potential bias in the estimated role of each mediator. We expect that these mediators will account for a substantial portion of the association between first-grade and age 15 mathematics achievement.

Method

Data

Data for our study came from the NICHD SECCYD. Babies born in designated 24-hr periods in 1991 in 10 locations around the United States were eligible to participate if their mothers were over 18 years old, spoke English as a first language, did not have any serious health conditions, did not plan to relinquish parental rights, did not plan to move in the next 3 years, and resided within an hour of a study site. The study sample was diverse both economically and ethnically, although it was not nationally representative. The recruitment process suffered from nearly 50% attrition, which was concentrated among low-socioeconomic-status mothers and children. More detail on recruitment procedures, data collection, and study procedures is provided by the NICHD Early Child Care Research Network (2002).

Reflecting the middle-class nature of the sample, the average family income-to-needs ratio measured repeatedly between birth and 1 month of age was 3.5. Furthermore, 76% of the sample identified as non-Hispanic White, 13% as African American, and study mothers reported an average of 14 years of education. More details regarding participant background can be found in the Supporting Information available online.

The current study used a subsample of children from the SECCYD (n = 1,362). Among the 1,364 children in the full sample, 2 had been in both special education and gifted and talented programs between second and fourth grades and were excluded from this study. To account for missing data, multiple imputation was used. All variables were included in the imputation models, and we also included a set of auxiliary variables to help ensure unbiased estimation. A total of 25 data sets were imputed using the ICE procedure in Stata 13.1. Stata MI commands were used for model estimation and statistical tests. Regression estimates were combined using Rubin's rules of combination (Rubin, 1987).

Measures

Academic Achievement in First Grade and at Age 15

Children were administered the Woodcock–Johnson (WJ–R) Psycho-educational Battery–Revised (Woodcock, McGrew, & Mather, 2001) as a measure of academic achievement at multiple times throughout the study. The current study focused on first-grade and age 15 scores on the Applied Problems subtest of the WJ–R. This commonly used measure of mathematics ability assessed knowledge of a wide range of mathematical topics and took about 15 min to complete. Students progressed through the examination until they incorrectly answered six consecutive questions. In first grade, most questions concerned principles of counting and simple addition and subtraction; at age 15, questions concerned algebraic concepts and procedures, among other topics.

We also included a measure of first-grade reading achievement in the analysis, as a covariate. Our measure of first-grade reading achievement involved questions taken from three components of the WJ–R: Word Attack, Letter–Word Identification, and Picture Vocabulary. Word Attack measured ability to read unfamiliar words using phonic and structural analysis, Letter–Word Identification asked students to identify and read aloud isolated letters and words, and Picture Vocabulary measured recognition and naming of familiar objects.

For all subtests taken from the WJ–R, we used *W* scores, which were centered at a mean of 500. This value represents the average score for a fifth-grade student. We then standardized the *W* scores based on our sample's mean and standard deviation. Testing procedures, reliability, and validity of the WJ–R subtests are widely available.

Mediators

Our mediators were drawn from child assessments taken between second and sixth grades. Although we would have preferred for all mediators to have been measured in the same waves, the pattern of data collection in the study dictated that we draw measures from several different waves. Key mediators of the associations between first-grade and age 15 mathematics achievement include fractions and division knowledge in fifth grade, EF in third and fourth grades, academic SCA in sixth grade, and school placements between second and fourth grades.

Knowledge of Fractions and Division

Our measure of division and fractions knowledge was taken from the WJ–R Calculation subtest, which was administered at fifth grade. The Calculation subtest assessed arithmetic and computational skills. We used the proportion of correct responses to Items 16, 22, 25, 26, 27, 28, 29, 30, 31, 38, and 42 of the test to derive our measure of division and fractions ability. We excluded questions that involved using division or fractions to solve algebra problems, as we did not want our measure of division and fractions knowledge to overlap conceptually with our age 15 measure of mathematics achievement.

Executive Functioning

EF was assessed across three separate measures in third and fourth grades. The first component of EF, working memory, was measured using the WJ–R Memory for Sentences task. The measure was administered in the laboratory during third grade; students were asked to listen to a tape player and repeat words, phrases, and sentences. As with the first-grade WJ–R subtests, we used the *W* score for the Memory for Sentences measure.

Third graders were also presented the Tower of Hanoi (TOH) task as a broad assessment of their EF skills. Research has indicated that the TOH task assesses planning and problem solving, working memory functioning, and inhibition (see Bull, Espy, & Senn, 2004). Children were directed to plan a sequence of moves to transport a set of disks stacked on one of three pegs, ordered from smallest to largest, to another peg, without ever placing a larger disk on a smaller one. The TOH included six increasingly difficult problems. Students were given an efficiency score based on how many trials they needed to complete each task (Borys, Spitz, & Dorans, 1982). The current analysis used the total number of moves across the six trials. The TOH is a well-validated and commonly used measure of EF skills in developmental psychology.

To measure sustained attention and inhibitory control, fourth graders were presented the Continuous Performance Task (CPT; Rosvold, Mirsky, Sarason, Bransome, & Beck, 1956). The task involved viewing pictures of common objects on a computer screen and pressing a button each time a target stimulus appeared. This task was repeated across 10 trials. We used the proportion of correct responses to the target stimulus as a measure of attention regulation, a common use of the CPT (e.g., Duncan et al., 2007).

Self-Concept

Our SCA measures were taken from the How I Do in School Scale, administered in a laboratory during sixth grade. The SCA items were adapted from the Self and Task Perception Questionnaire (Jacobs, Lanza, Osgood, Eccles, & Wigfield, 2002); on them, students were asked to indicate their beliefs regarding their abilities in mathematics (five items). Students responded to questions such as "How good at math are you?" using a 7-point Likert scale ranging from 1 = *not at all* to 7 = *very*. Students were also administered similar SCA items designed to measure English self-perceptions (five items), which we used in order to test the domain specificity of math SCA. The composite scores for math and English were calculated by taking the mean across the five items for each scale. The math and English self-concept scales have reported alphas of .82.

School Placement

Selection into special education or gifted and talented programs was measured through a teacher report survey at Grades 2, 3, and 4. Teachers were asked to report how many hours per week the child was currently receiving special education or gifted and talented services. Both the gifted and talented measure and the special education measure were coded 1 if the student received any such services in any of the three grades and 0 otherwise.

Additional Covariates

To reduce bias in the estimated effects of first-grade academic achievement on subsequent EF and age 15 achievement, we controlled for first-grade EF, as measured by the WJ–R Memory for Sentences test, the TOH, and the CPT. To account for bias in our estimates due to correlations between achievement and behavioral problems, we also controlled for first-grade measures of inattentiveness, aggressive behavior, and internalizing behavior, as assessed by the Teacher Report Form, an instrument based on the Child Behavior Checklist. We also included an assessment of child social skills, measured by teacher report on the Social Skills Rating System (Gresham & Elliott, 1990).

Controls taken prior to first grade included scores on the Surgency, Negative Affectivity, and Effortful Control subscales of the Child Behavior Questionnaire. These measures of temperament reflected mothers' reports at the 54-month child interview. We controlled for early cognitive functioning using the Bracken Basic Concept Scale School Readiness score taken at 36 months, and the Bayley Mental Development Index taken at 24 months.

Child and Family Demographic Characteristics

To account for child and family characteristics that could be correlated with achievement and our proposed mediators, we included a host of child and family background covariates. Child demographic controls included dummy variables for the sites at which the family was recruited to participate in the study, birth weight in grams; gender (1 = boy), race (non-Hispanic African American, Hispanic, or other, with non-Hispanic White as the reference group), and age in years at the time of first-grade testing. Family background characteristics included the mean number of children in the household between the child's 24-month and first-grade assessments, the natural logarithm of family income averaged between the 1-month and first-grade assessments, the education of the mother in years at the time of the 1-month assessment, and mother's age at the birth of the target child.

Results

Descriptive statistics for key variables are shown in Table 1 (for descriptive statistics for all analysis variables, including controls, see Table S1 available in the online Supporting Information). Recall that our measures of student achievement were WJ–R *W* scores centered at a value of 500 (the average score for a typically achieving fifth grader). Participants' growth between first grade and age 15 in both mathematics and reading is shown in Table 1. Not surprisingly, students scored well below 500 in first grade, and well above this score at age 15.

Table 1 also presents descriptive statistics for our proposed mediators. Participants scored rather high on the SCA scale for mathematics, as the average score on this scale (ranging from 1 to 7, with 7 indicating very high mathematics SCA) was 5.76 (SD = 1.01). Between second and fourth grades, 10% of our sample had been designated for special education services and 15% had been recommended for a gifted and talented program. Indeed, our data differ from current estimates of the percentage of students placed in such programs (13% for special education services and 7% for gifted and talented programs; Snyder & Dillow, 2013). Thus, our results should not be generalized to the national level, as our data are not nationally representative. Among our EF measures, the average proportion of correct responses to target stimuli on the CPT task was high (M = 0.95), and participants in our sample who took the WJ–R working memory measure during third grade scored near the Grade 5 average (M = 494, SD = 14). Finally, on average, participants correctly answered 44% of the fraction and division questions presented on the WJ–R Calculation subtest.

Zero-order correlations among the measures of academic achievement and the mediators taken between first grade and age 15, inclusive, are presented in Table 2 (correlations between key variables and all analysis variables can be found in the online Supporting Information). As might be expected, correlations among the various WJ achievement subscales were the highest, ranging from .48 to .74. Among our mediators, fifth-grade fraction and division knowledge had the highest correlation with age 15 mathematics achievement,

Table 1
Descriptive Statistics for Primary Analysis Variables

	M	*SD*	Min	Max
Age 15 achievement				
Mathematics	524	16	444	591
First-grade achievement				
Mathematics	470	15	408	516
Reading	470	15	417	512
Mediators:				
Math SCA (sixth grade)	5.76	1.01	1	7
School placements (second to fourth grades)				
Special education	0.10	0.30	0	1
Gifted and talented	0.15	0.35	0	1
Executive functions (third to fourth grades)				
Memory for Sentences	494	14	409	539
Tower of Hanoi	17	8	0	35
Continuous Performance Task	0.95	0.07	0.36	1
Division and fraction knowledge (fifth grade)	0.44	0.22	0	1
First-grade executive function controls				
Memory for Sentences	481	15	368	529
Tower of Hanoi	14	7	0	34
Continuous Performance Task	0.95	0.07	0.41	1

Note. Descriptive statistics generated from 25 multiply imputed data sets (n = 1,362 each). Achievement measures, mathematics and reading, and Memory for Sentences (working memory) were drawn from Woodcock–Johnson (WJ–R) *W* scores. Math self-concept of ability (SCA) is scaled from 1 to 7, with a score of 7 indicating more confidence in one's ability. School placements were measured by the proportion of years the student was placed in special education or gifted and talented services between second and fourth grade. The Tower of Hanoi is considered to measure planning and a broad array of executive functioning skills. The planning score indicates the total number of moves a child needed to complete the task across six trials. Continuous Performance Task is measured by calculating the proportion of correct responses to target stimuli, and it is understood to be a measure of attention regulation. Division and fraction knowledge was calculated as a proportion of items correctly answered on the WJ–R Calculation subtest.

Table 2
Correlations Between Mathematics Achievement and Mediators

	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Math (age 15)	1												
2. Math (first grade)	.59	1											
3. Reading (first grade)	.50	.60	1										
4. Division and fraction skills (fifth grade)	.54	.50	.47	1									
5. Math self-concept of ability (sixth grade)	.34	.26	.11	.29	1								
6. Special education (second to fourth grades)	−.27	−.30	−.36	−.27	−.10	1							
7. Gifted and talented (second to fourth grades)	.26	.27	.29	.21	.06	−.14	1						
8. Memory for Sentences (third grade)	.48	.50	.50	.32	.15	−.28	.22	1					
9. Tower of Hanoi (third grade)	.29	.34	.23	.29	.18	−.19	.14	.25	1				
10. Continuous Performance Task (fourth grade)	.26	.25	.24	.22	.09	−.17	.10	.24	.22	1			
11. Memory for Sentences (first grade)	.45	.50	.50	.28	.12	−.23	.24	.74	.22	.18	1		
12. Tower of Hanoi (first grade)	.21	.25	.20	.18	.07	−.16	.12	.19	.38	.13	.18	1	
13. Continuous Performance Task (first grade)	.25	.21	.24	.20	.10	−.25	.10	.22	.18	.31	.14	.13	1

Note. All correlations presented in this table were obtained from 25 multiply imputed data sets and were statistically significant at the .05 level.

$r(1{,}362) = .54$, $p < .05$, followed by proficiency on the Memory for Sentences task in third grade, $r(1{,}362) = .48$, $p < .05$. Interestingly, mathematics SCA and English SCA were positively correlated, $r(1{,}362) = .28$, $p < .05$ (not shown in Table 2); high self-perception in one domain did not predict a lower self-perception in the other. Finally, many mediators had moderate to high correlations with one another and with age 15 achievement. This demonstrates the bias-reducing utility of accounting for multiple intermediary pathways when engaging in mediational analyses.

Tables 3 and 4 summarize results from ordinary least squares (OLS) and logistic regression analyses in which mediational hypotheses were tested in two steps. First, we tested the relation between each mediator and first-grade mathematics achievement, controlling for other cognitive abilities and background characteristics. Next, we related age 15 mathematics achievement to first-grade mathematical skills and controls. We then introduced our set of mediators into the model. We present alternative analytic approaches, including results estimated via path analysis, in Tables S2 and S3. We also describe attempts to estimate our data via structural equation modeling (SEM) and the analytical problems that led us to prefer the estimates presented here. In brief, we found no evidence that our EF measures should be combined as a single latent factor. Without latent factors, any SEM analysis will directly replicate the OLS results. Furthermore, the SEM approach introduced problems with missing data, so we chose OLS as our preferred estimation strategy.

Predicting the Mediators

Table 3 displays regression results from a series of OLS and logistic regression models in which each mediator was used as the dependent variable and first-grade mathematics achievement was the key independent variable. The results shown are from fully controlled models that included reading achievement, other first-grade characteristics, and measures of family background (see table note for full list of covariates). We used OLS regression to relate third-grade EF, fifth-grade fractions and division knowledge, and sixth-grade SCA to first-grade mathematics achievement and covariates. A corresponding multinomial logistic model regressed special education and gifted and talented placements on the same predictors.

First-grade mathematics achievement was a significant predictor of each of the mediators (Table 3). First-grade mathematics achievement strongly predicted later math SCA ($\beta = .29$, $SE = 0.04$, $p < .001$), and first-grade reading was not a significant predictor of math SCA. Thus, we did not find evidence of an opposing relation between achievement in reading and math SCA. As a further robustness check to examine the domain specificity of SCA, we also related sixth-grade English SCA to first-grade mathematics achievement and covariates. In results not shown in Table 3, mathematics achievement was not a significant predictor of later English SCA, but first-grade reading achievement was ($\beta = .11$, $SE = 0.05$, $p < .01$). These results suggest that children develop domain-specific SCAs, but in our data, high achievement in one school subject did not affect SCA in other areas.

Table 3
Ordinary Least Squares (OLS)/Multinomial Logistic (MNL) Regression Estimates of the Association Between First-Grade Mathematics Achievement and Analysis Mediators

Dependent variables	Coefficients and standard errors for first-grade math β	*SE*
OLS regression coefficients and standard errors		
Self-concept		
Math self-concept of ability (SCA)	.29***	(0.04)
Reading self-concept of ability (SCA)	−.06	(0.05)
Executive functions		
Memory for Sentences	.10*	(0.05)
Tower of Hanoi	.19***	(0.04)
Continuous Performance Task	.12	(0.06)
Division and fraction knowledge	.31***	(0.04)
Multinomial logistic regression coefficients and standard errors		
School placements (MNL)		
Special education	−.54*	(0.22)
Gifted and talented	.38*	(0.15)

Note. Dependent variables are listed on the left (rows). All coefficients listed were produced from fully controlled models. All continuous variables were standardized to *z* scores using the means and standard deviations from each imputed data set (25 multiply imputed data sets; $n = 1{,}362$). The likelihood of placement into special education and gifted and talented programs was estimated using multinomial logistic regression, with placement into no program as the comparison group. Standard errors were obtained by robust estimator. Covariates included first-grade measures of reading achievement, working memory, planning, sustained attention, externalizing, internalizing, and social skills; 54-month measures of surgency, negative affectivity, and effortful control; a 36-month measure of school readiness (Bracken Basic Concept Scale) and a 24-month measure of mental development (Bayley Mental Development Index); birth weight, ethnicity, gender, age at first-grade test, number of children in the home, average family income between 1 month and first grade, mother's education level, mother's age at birth, and site.
*$p \le .05$. ***$p \le .001$.

First graders' math knowledge strongly predicted EF skills, including performance on the working memory task (β = .10, *SE* = 0.05, $p < .05$) and on the TOH (β = .19, *SE* = 0.04, $p < .001$). It negatively predicted special education placement (β = −.54, *SE* = 0.22, $p < .05$), and positively predicted gifted and talented placement (β = .38, *SE* = 0.15, $p < .05$) and fifth-grade fraction and division knowledge (β = .31, *SE* = 0.04, $p < .001$).

Predicting Age 15 Achievement

Regression results from a series of models estimating the association between first-grade mathematics and age 15 achievement are shown in Table 4. All models include the extensive set of controls listed at the bottom of the table. Coefficients on these covariates can be found in Table S4.

Results in Model 1 of Table 4 show that after adjusting for our extensive set of baseline controls, a 1 *SD* increase in first-grade mathematics was associated with a 0.36 *SD* (*SE* = 0.05, $p < .001$) increase in age 15 math achievement. First-grade measures of reading (β = .10, *SE* = 0.04, $p < .05$), working memory (β = .10, *SE* = 0.04, $p < .05$), and CPT (β = .08, *SE* = 0.04, $p < .05$) were significant, but weaker, predictors of age 15 math achievement. These coefficients are remarkably similar to the averages reported in the six study synthesis of school-readiness skills in Duncan et al. (2007), who examined the association between a similar set of predictors and late elementary school achievement.

Model 2 displays results of age 15 mathematics achievement models that include three of our four mediators: EF, SCA, and school placements. A comparison of coefficients on the early achievement measures between Models 1 and 2 showed that these mediators accounted for about 25% of the early to later mathematics associations and none of the early reading to later math associations. The .09 reduction in the math coefficient was statistically significant ($p < .01$). Two of these mediators had significant associations with math outcomes: math SCA (β = .20, *SE* = 0.02, $p < .001$) and placement into gifted and talented programs (β = .20, *SE* = 0.07, $p < .01$).

To further test our hypothesis that SCA contributes to achievement through domain-specific channels, we also tested a mediational model (not shown in Table 4) in which age 15 mathematics achievement was related to first-grade characteristics, all four mediators, and English SCA. English SCA failed to produce a statistically significant coefficient, and it did not diminish the coefficient produced by math SCA (β = .20, *SE* = 0.02, $p < .001$).

Model 3 of Table 4 adds Grade 5 fraction and division knowledge to the other three mediators and shows that it also significantly predicted later mathematics achievement. Moreover, when all of our mediators were considered in one model, 39% of the early-to-later mathematics effect was accounted for. This amounted to a .13 reduction in the first-grade mathematics achievement coefficient ($p < .001$; see Model 4).

Contrary to expectations, none of the three components of third - or fourth - grade EF was predictive of later math achievement. Note that our set of control variables includes first-grade EF. Substantial correlations (*r*s = .25 to .50) were present between third- and fourth-grade EF and age 15 mathematics score, but the corresponding correlations between

Table 4
Regression Adjusted Estimates of the Associations Between First-Grade Achievement and Age 15 Mathematics Achievement

	Model 1		Model 2		Model 3		Model 4	
	Age 15 math (baseline)		Age 15 math (mediators)		Age 15 math (mediators + fraction and division knowledge)		Coefficient reduction from Model 1 to Model 3	
	β	*SE*	β	*SE*	β	*SE*	β	*SE*
First-grade achievement								
Mathematics	.36***	(0.05)	.27***	(0.05)	.22***	(0.05)	.13***	(0.02)
Reading	.10*	(0.04)	.10*	(0.04)	.05	(0.04)		
Mediators:								
Math self-concept of ability			.20***	(0.02)	.16***	(0.02)		
School placements (second to fourth grades)								
Special education			−.13	(0.09)	−.10	(0.09)		
Gifted and talented			.20**	(0.07)	.17*	(0.07)		
Executive function (third to fourth grades)								
Memory for Sentences			.08	(0.05)	.08	(0.05)		
Tower of Hanoi			.03	(0.03)	.01	(0.03)		
Continuous Performance Task			.04	(0.04)	.03	(0.03)		
Division and fractions knowledge (fifth grades)					.21***	(0.03)		
First-grade executive function								
Memory for Sentences	.10*	(0.04)	.05	(0.06)	.07	(0.06)		
Tower of Hanoi	.03	(0.03)	.01	(0.03)	.01	(0.03)		
Continuous Performance Task	.08*	(0.04)	.04	(0.03)	.04	(0.03)		
Covariates	Inc.		Inc.		Inc.			

Note. Continuous variables were standardized to *z* scores using the means and standard deviations from each imputed data set (25 multiply imputed data sets; $n = 1{,}362$). Standard errors were obtained by the robust estimator. All coefficients were generated from fully controlled models. Covariates included first-grade measures of externalizing, internalizing, and social skills; 54-month measures of surgency, negative affectivity, and effortful control; a 36-month measure of school readiness (Bracken Basic Concept Scale) and a 24-month measure of mental development (Bayley Mental Development Index); birth weight, ethnicity, gender, age at first-grade test, number of children in the home, average family income between 1 month and first grade, mother's education level, mother's age at birth, and site. Model 4 was obtained by using seemingly unrelated estimation.
$^{*}p \le .05$. $^{**}p \le .01$. $^{***}p \le .001$.

first-grade EF and age 15 mathematics were almost as high (Table 2). When both were included in the regressions, the associations between age 15 mathematics and third- and fourth-grade EF were no stronger than the associations between age 15 mathematics and first-grade EF.

We also estimated models using an aggregated measure of EF, which averaged the three measures of EF into one variable. These models did not produce a substantially different mediational effect; third- and fourth-grade EF was still an insignificant predictor of age 15 mathematics. Models with the aggregated measure of EF are presented in Table S5.

Discussion

The current study tested for mediators of the strong relation between early and later mathematics achievement. We found that the relation was mediated by fractions and division knowledge, math self-concept, and placement into gifted and talented programs. Together, these mediators accounted for over one third of the association between mathematics achievement in first grade and at age 15.

Our study was conducted within a bioecological framework (Bronfenbrenner & Morris, 2006), as we hypothesized that a combination of environmental and personal characteristics would shape long-run mathematics achievement trajectories. Indeed, we found evidence that both environmental (placement into gifted and talented programs) and personal (mathematics SCA and fraction and division knowledge) factors influenced long-run achievement in mathematics.

Although cognitive skill-building theories have received much recent attention in regard to the development of academic achievement (e.g., Cunha

& Heckman, 2008; Siegler, Fazio, Bailey, & Zhou, 2013), our results suggest that factors beyond mathematical skill attainment play a role in the development of mathematics achievement. Even when controlling for fraction and division knowledge in fifth grade, which appears to be crucial for later mathematics, mathematics SCA and participation in gifted and talented programs were still strong predictors of later achievement. Thus, although earlier mathematical skills are critical for obtaining later ones, the influence of other factors, including motivation and school context, should not be overlooked.

Our findings regarding SCA supported our hypothesis that early mathematics achievement would lead to the development of positive SCA in mathematics, which would in turn support students' later math achievement. It has been theorized that SCA may bolster subsequent achievement in mathematics by altering student motivation and mathematics-related affect, thereby increasing the likelihood of engaging in mathematics and spending time studying (e.g., Bong & Skaalvik, 2003). Although our results appear to support this hypothesis, we cannot rule out other avenues through which SCA may contribute to subsequent achievement (e.g., parental or teacher influence).

Our findings also give some indication of how self-perceptions in mathematics arise. The finding that early mathematics achievement predicted mathematics SCA in sixth grade, after controlling for other relevant variables, suggests that students assess their own mathematics abilities with some accuracy. We also found evidence for the domain specificity of SCA, as early mathematics achievement predicted subsequent math-specific self-concepts, but not subsequent English self-perceptions. However, we found a positive correlation between mathematics SCA and English SCA in Grade 6. Some have argued that SCA in math and English should have a null, or even negative, correlation (e.g., Marsh et al., 1988), as children may view themselves as "good" in one subject, but less equipped in others. Yet, mathematics and reading achievement are highly correlated, and if students only saw themselves as proficient in one subject, they would have an inaccurate perception of their abilities in the other subject.

The current findings suggest that this was not the case. First, we found that previous achievement in each domain best predicted later SCA. Second, previous achievement in mathematics, controlling for reading, had no relation to later reading SCA, and the same was true of early reading and later math SCA. Thus, although we found evidence of domain specificity, we did not find an opposing relation between SCA in mathematics and English.

We also found that placement into gifted and talented programs positively predicted later achievement, even after controlling for a host of background variables and concurrent mediators. This finding suggests the value of such programs at a time that funding reductions threaten their continued existence (Fleming, 2013). Importantly, experimental research is needed to estimate the causal connections between participation in gifted and talented programs and student achievement, especially because selection into such programs could bias the associations reported by studies such as this one. Nonetheless, reducing or eliminating funding without evidence from experimental studies seems premature in light of the positive effects found here and in quasi-experimental work (e.g., Bhatt, 2009; but see Bui et al., 2011). Relatedly, future work should attempt to uncover the mechanisms through which such programs support achievement development, as the "gifted and talented" label undoubtedly describes a heterogeneous set of programs. Improving our understanding of which types of gifted and talented programs are most successful can further our ability to tailor services for these students.

Unsurprisingly, we found that early mathematics achievement negatively predicted placement into special education. Indeed, schools have increasingly relied upon measures of early academic skills for diagnosing students at risk for underachievement, as response to intervention and multitiered systems of instruction have become prevalent across the country (see Fuchs, Fuchs, & Stecker, 2010). Our results provide some support for this practice, as first-grade mathematics ability strongly predicted high school mathematics achievement. However, our study did not find any relation between placement into special education services and later achievement. This could be seen as evidence supporting efforts to "mainstream" children out of special education programs, but null effects on student achievement in a correlational model could also be a sign of the programs' successes. Depending on each student's learning disadvantages, special education programs may be effective if they merely prevent students from falling further behind. Thus, estimates from correlational studies could contain downward bias, as students placed into special education are more likely to experience negative growth in achievement compared to other students. If this were the case, and special education programs did nothing to curb this trend, then we would expect a negative association between placement into services and later achievement. As with gifted and talented programs, this issue further illuminates the strong need for experimental research in order to draw clear conclusions regarding the effectiveness of these services.

Our findings regarding the lack of mediation for the EF measures converge with recent experimental research. Taken together, this work indicates that attempting to change specific aspects of EF with the hope of affecting mathematics achievement may not be a useful strategy. Experimental evaluations of Tools of the Mind, a curriculum meant to boost young children's EF, found that the curriculum failed to change EF skills or mathematics achievement (Clements et al., 2015; Wilson & Farran, 2012). Clements and colleagues (2015) did find some indication that experimentally induced boosts in early mathematics achievement may cause later improvements in EF. Although the possibility of boosting EF through early mathematics intervention may be theoretically interesting, most educational interventions hope to change academic skills. Our results imply that improved EF skills do not readily transfer to improved mathematics ability.

The present results also suggest that efforts to improve math achievement through EF interventions in later elementary school would probably have little effect. Our results likely represent upper bound estimates of the causal effect of each of our respective EF measures on later achievement, as any bias due to omitted variables is likely positive; plausible confounding variables (e.g., parental support for cognitive development, interest in mathematics) are probably positively correlated with the EF measures and adolescent mathematics achievement. However, measurement error could bias the effect of EF downward, and if EF describes a single latent construct, there is little reason to expect that our three measures perfectly captured it. Nonetheless, we employed well-validated measures to assess various aspects of EF (e.g., Bull et al., 2004; Duncan et al., 2007; Watts et al., 2014). Although the TOH has been criticized for its poor test–retest reliability (estimates range from .50 to .75; see Beck, Schaefer, Pang, & Carlson, 2011), there was no indication that the other two, more reliable, measures of working memory (Memory for Sentences) and sustained attention (CPT) positively affected later mathematics achievement.

On the other hand, our findings illustrate the importance of domain-specific knowledge in the development of long-term mathematics achievement, as knowledge of fractions and division substantially mediated long-run mathematics achievement associations. This finding converges with results from prior studies (e.g., Beier & Ackerman, 2005) that have demonstrated the relative importance of content knowledge over domain-general cognitive competencies. Thus, finding crucial areas of content knowledge that strongly relate to subsequent mathematics achievement may provide a fruitful means through which educational practitioners can intervene to improve long-term achievement trajectories. Indeed, practitioners and researchers have designed interventions and instructional practices that improve content knowledge of fractions in late elementary and early middle school and that transfer to overall mathematics achievement (for a review, see Siegler et al., 2013).

Although our set of mediators had some success in accounting for the association between first-grade and age 15 mathematics achievement, the majority of this association remained unexplained in our models. We certainly did not consider every possible mediator, as we were limited by the measures included in the data. In particular, classroom instructional practices and classroom climate were not measured in our study. Although certain instructional practices, such as emphasizing advanced content (Claessens, Engel, & Curran, 2014; Engel, Claessens, & Finch, 2013), have been linked to elementary school mathematics gains, any mediational analyses covering a time span of multiple years would need to consider instruction in each year. Further complicating matters, many students have multiple teachers in any given year. Unfortunately, analyses of such factors were not possible with our data.

It could also be argued that the current study was limited by the narrow time window during which each mediator was measured. With the exception of school placements, our mediators were only assessed at one point between first grade and age 15, and the mediators were not assessed at the same time. Future research should examine mediators measured at multiple time points in order to understand the longitudinal pathways that may be occurring. At the same time, previous research has indicated that individual differences in both SCA (Davis-Kean et al., 2008) and EF (Mazzocco & Kover, 2007) are quite stable throughout elementary school. Thus, although we cannot rule out the possibility, we doubt that our results would differ substantially if our measurements of these constructs had been obtained in multiple grades.

Additionally, our measures of school placements were taken over the course of 3 years, yet we only measured whether a student was ever placed in a special education or gifted and talented program during this time span. Future research should investigate possible dosage effects of school placements on mathematics achievement, as spending more time in one of these programs could have a compounding effect on mathematical development.

Our inability to fully account for the early to later mathematics association may also indicate that some of this association is not causal. Indeed, recent work by Bailey, Watts, Littlefield, and Geary (2014) suggests that even highly controlled longitudinal examinations of mathematics achievement fail to account for all of the underlying personal and environmental characteristics that contribute to achievement trajectories. Consequently, such analyses may overestimate the causal effect of early achievement on later achievement. Moreover, recent experimental work has raised questions regarding the causal association between early and later measures of achievement. Clements, Sarama, Wolfe, and Spitler (2013) found that a preschool mathematics intervention produced a large impact on end-of-preschool mathematics achievement, but this effect faded by over 60% by the end of first grade. Thus, although highly controlled, correlational studies (e.g., Duncan et al., 2007; Watts et al., 2014) have found strong associations between early and later measures of achievement, it remains unclear whether intervention-produced gains in early knowledge produce similar effects.

A final implication of these findings is that experimental studies with longitudinal follow-ups are greatly needed. We found that individual differences in mathematics achievement between first grade and high school are quite stable, and we identified several processes, acting at both the personal and environmental levels, that contribute to stability and change in mathematics achievement. Although our models shed some light on how these long-run trajectories develop, only experimental research can determine whether these trajectories can be meaningfully altered by interventions. Our results may indicate possible areas in which future experimental work could be conducted. However, our study illustrates that a number of processes (e.g., self-concept, school placements, specific math skills) appear to simultaneously influence mathematics achievement development. Thus, interventions should be conducted with realistic expectations, as they are likely to be unsuccessful if they fail to take into account the diverse processes involved in children's mathematical development and in their cognitive development more generally.

References

Bailey, D. H., Siegler, R. S., & Geary, D. C. (2014). Early predictors of middle school fraction knowledge. *Developmental Science*, *17*, 775–785. doi:10.1111/desc.12155

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

Beck, D. M., Schaefer, C., Pang, K., & Carlson, S. M. (2011). Executive function in preschool children: Test–retest reliability. *Journal of Cognition and Development*, *12*, 169–193. doi:10.1080/15248372.2011.563485

Beier, M. E., & Ackerman, P. L. (2005). Age, ability and the role of prior knowledge on the acquisition of new domain knowledge. *Psychology and Aging*, *20*, 341–355. doi:10.1037/0882-7974.20.2.341

Bhatt, R. R. (2009). *The impacts of gifted and talented education* (Andrew Young School of Policy Studies Research Paper Series No. 09-11). Retrieved from http://ssrn.com/abstract=1494334

Blackorby, J., & Wagner, M. (1996). Longitudinal post-school outcomes of youth with disabilities: Findings from the National Longitudinal Transition Study. *Exceptional Children*, *26*, 399–413. doi:10.1177/001440299606200502

Blair, C., & Razza, R. P. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Development*, *78*, 647–663. doi:10.1111/j.1467-8624.2007.01019.x

Bong, M., & Skaalvik, E. M. (2003). Academic self-concept and self-efficacy: How different are they really? *Educational Psychology Review*, *15*, 1–40. doi:10.1023/A:1021302408382

Booth, J. L., & Newton, K. J. (2012). Fractions: Could they really be the gatekeeper's doorman? *Contemporary Educational Psychology*, *37*, 247–253. doi:10.1016/j.cedpsych.2012.07.001

Borys, S. V., Spitz, H. H., & Dorans, B. A. (1982). Tower of Hanoi performance of retarded young adults and nonretarded children as a function of solution length and goal state. *Journal of Experimental Child Psychology*, *33*, 87–110. doi:10.1016/0022-0965(82)90008-X

Bronfenbrenner, U., & Morris, P. (2006). The bioecological model of human development. In W. Damon & R. M. Lerner (Eds.), *Handbook of child psychology: Theoretical models of human development* (Vol. 1, 6th ed., pp. 793–828). New York, NY: Wiley.

Bui, S. A., Craig, S. G., & Imberman, S. A. (2011). *Is gifted education a bright idea? Assessing the impact of gifted and talented programs on achievement* (Working Paper No. w17089). Cambridge, MA: National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w17089

Bull, R., Espy, K. A., & Senn, T. E. (2004). A comparison of performance on the Towers of London and Hanoi in young children. *Journal of Child Psychology and Psychiatry*, *45*, 743–754. doi:10.1111/j.1469-7610.2004.00268.x

Claessens, A., & Engel, M. (2013). How important is where you start? Early mathematics knowledge and later school success. *Teachers College Record*,*13*, 1–29. Retrieved from http://www.tcrecord.org. ID Number: 16980

Claessens, A., Engel, M., & Curran, F. C. (2014). Academic content, student learning, and the persistence of preschool effects. *American Educational Research Journal*, *51*, 403–434. doi:10.3102/0002831213513634

Clark, C. A., Pritchard, V. E., & Woodward, L. J. (2010). Preschool executive functioning abilities predict early mathematics achievement. *Developmental Psychology*, *46*, 1176. doi:10.1037/a0019672

Clements, D. H., Sarama, J., Layzer, C., Unlu, F., Germeroth, C., & Fesler, L. (2015). *Effects on executive function and mathematics learning of an early mathematics curriculum synthesized with scaffolded play designed to promote self-regulation versus the mathematics curriculum alone*. Manuscript submitted for publication.

Clements, D. H., Sarama, J., Wolfe, C. B., & Spitler, M. E. (2013). Longitudinal evaluation of a scale-up model for teaching mathematics with trajectories and technologies: Persistence of effects in the third year. *American Educational Research Journal*, *50*, 812–850. doi:10.3102/0002831212469270

Cook, B. G., & Schirmer, B. R. (2003). What is special about special education? Overview and analysis. *Journal of Special Education*, *37*, 200–205. doi:10.1177/00224669030370031001

Cunha, F., & Heckman, J. J. (2008). Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation. *Journal of Human Resources*, *43*, 738–782. doi:10.3368/jhr.43.4.738

Davis-Kean, P. E., Huesmann, L. R., Jager, J., Collins, W. A., Bates, J. E., & Lansford, J. E. (2008). Changes in the relation of self-efficacy beliefs and behaviors across development. *Child Development*, *79*, 1257–1269. doi:10.1111/j.1467-8624.2008.01187.x

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., . . . Japel, C. (2007). School readiness and later achievement. *Developmental Psychology*, *43*, 1428–1446. doi: 10.1037/0012-1649.43.6.1428

Engel, M., Claessens, A., & Finch, M. A. (2013). Teaching students what they already know? The (mis) alignment between mathematics instructional content and student knowledge in kindergarten. *Educational Evaluation and Policy Analysis*, *35*, 157–178. doi:10.3102/0162373712461850

Fleming, N. (2013, October). Parents press for attention to programs for gifted students. *Education Week*. Retrieved from http://www.edweek.org/ew/articles/2013/10/02/06giftedparents_ep.h33.html

Fuchs, D., & Fuchs, L. S. (1998). Researchers and teachers working together to adapt instruction for diverse learners. *Learning Disabilities Research & Practice*, *13*, 126–137.

Fuchs, D., Fuchs, L. S., & Stecker, P. M. (2010). The "blurring" of special education in a new continuum of general education placements and services. *Exceptional Children*, *76*, 301–323. doi:10.1177/001440291007600304

Fuhs, M. W., Nesbitt, K. T., Farran, D. C., & Dong, N. (2014). Longitudinal associations between executive functioning and academic skills across content areas. *Developmental Psychology*, *50*, 1698–1709. doi:10.1037/a0036633

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

Geary, D. C., Hoard, M. K., Nugent, L., & Bailey, D. H. (2013). Adolescents' functional numeracy is predicted by their school entry number system knowledge. *PLoS ONE*, *8*, e54651. doi:10.1371/journal.pone.0054651

Gresham, F. M., & Elliott, S. N. (1990). *Social Skills Rating System (Elementary Scale A)*. Circle Pines, MN: American Guidance Service.

Hanushek, E. A., Kain, J. F., & Rivkin, S. G. (2002). Inferring program effects for special populations: Does special education raise achievement for students with disabilities? *Review of Economics and Statistics*, *84*, 584–599. doi:10.1162/003465302760556431

Helmke, A., & van Aken, M. A. (1995). The causal ordering of academic achievement and self-concept of ability during elementary school: A longitudinal study. *Journal of Educational Psychology*, *87*, 624–637. doi:10.1037/0022-0663.87.4.624

Hibel, J., Farkas, G., & Morgan, P. L. (2010). Who is placed into special education? *Sociology of Education*, *83*, 312–332. doi:10.1177/0038040710383518

Horn, W. F., & Tynan, D. (2001). Time to make special education "special" again. In C. E. Finn Jr., A. J. Rotherham, & C. R. Hokanson Jr. (Eds.), *Rethinking special education for a new century* (pp. 23–51). Washington, DC: Fordham Foundation.

Imberman, S. A., Kugler, A. D., & Sacerdote, B. I. (2012). Katrina's children: Evidence on the structure of peer effects from hurricane evacuees. *American Economic Review*, *102*, 2048–2082. Retrieved from http://www.jstor.org/stable/41724614

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*, 509–527. doi:10.1111/1467-8624.00421

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

Lane, K. L., Wehby, J. H., Little, M. A., & Cooley, C. (2005). Students educated in self-contained classrooms and self-contained schools: Part II—How do they progress over time? *Behavioral Disorders*, *30*, 363–374. Retrieved from: http://www.jstor.org/stable/23889849

Marsh, H. W., Byrne, B. M., & Shavelson, R. J. (1988). A multifaceted academic self-concept: Its hierarchical structure and its relation to academic achievement. *Journal of Educational Psychology*, *80*, 366–380. doi:10.1037/0022-0663.80.3.366

Marsh, H. W., Byrne, B. M., & Yeung, A. S. (1999). Causal ordering of academic self-concept and achievement:

Reanalysis of a pioneering study and revised recommendations. *Educational Psychologist*, *34*, 155–167. doi:10.1207/s15326985ep3403_2

Marsh, H. W., & Craven, R. G. (2006). Reciprocal effects of self-concept and performance from a multidimensional perspective: Beyond seductive pleasure and unidimensional perspectives. *Perspectives on Psychological Science*, *1*, 133–163. doi:10.1111/j.1745-6916.2006.00010.x

Marsh, H. W., & Yeung, A. S. (1997). Causal effects of academic self-concept on academic achievement: Structural equation models of longitudinal data. *Journal of Educational Psychology*, *89*, 41–54. doi:10.1037/0022-0663.89.1.41

Martin, W. G., Strutchens, M. E., & Elliott, P. C. (2007). *The learning of mathematics*. Reston, VA: National Council of Teachers of Mathematics.

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

McClain, M. C., & Pfeiffer, S. (2012). Identification of gifted students in the United States today: A look at state definitions, policies, and practices. *Journal of Applied School Psychology*, *28*, 59–88. doi:10.1080/15377903.2012.643757

Morgan, P. L., Frisco, M. L., Farkas, G., & Hibel, J. (2010). A propensity score matching analysis of the effects of special education services. *Journal of Special Education*, *43*, 236–254. doi:10.1177/0022466908323007

NICHD Early Child Care Research Network. (2002). Early childcare and children's development prior to school entry: Results from the NICHD Study of Early Child Care. *American Educational Research Journal*, *39*, 133–164. doi:10.3102/00028312039001133

Nunes, T., Bryant, P., Barros, R., & Sylva, K. (2012). The relative importance of two different mathematical abilities to mathematical achievement. *British Journal of Educational Psychology*, *82*, 136–156. doi:10.1111/j.2044-8279.2011.02033.x

Rosvold, H. E., Mirsky, A. F., Sarason, I., Bransome, E. D., Jr., & Beck, L. H. (1956). A continuous performance test of brain damage. *Journal of Consulting Psychology*, *20*, 343.

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. Hoboken, NJ: Wiley.

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*, 691–697. doi:10.1177/0956797612440101

Siegler, R. S., Fazio, L. K., Bailey, D. H., & Zhou, X. (2013). Fractions: The new frontier for theories of numerical development. *Trends in Cognitive Sciences*, *17*, 13–19. doi:10.1016/j.tics.2012.11.004

Siegler, R. S., Thompson, C. A., & Schneider, M. (2011). An integrated theory of whole number and fractions development. *Cognitive Psychology*, *62*, 273–296. doi:10.1016/j.cogpsych.2011.03.001

Skaalvik, E. M., & Valås, H. (1999). Relations among achievement, self-concept, and motivation in mathematics and language arts: A longitudinal study. *Journal of Experimental Education*, *67*, 135–149. doi:10.1080/00220979909598349

Snyder, T., & Dillow, S. (2013). *Digest of education statistics, 2012*. Washington, DC: National Center for Education Statistics.

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

Vaughn, V. L., Feldhusen, J. F., & Asher, J. W. (1991). Meta-analyses and review of research on pull-out programs in gifted education. *Gifted Child Quarterly*, *35*, 92–98.

Watts, T. W., Duncan, G. J., Siegler, R. S., & Davis-Kean, P. E. (2014). What's past is prologue: Relations between early mathematics knowledge and high school achievement. *Educational Researcher*, *43*, 352–360. doi:10.3102/0013189X14553660

Welsh, J. A., Nix, R. L., Blair, C., Bierman, K. L., & Nelson, K. E. (2010). The development of cognitive skills and gains in academic school readiness for children from low-income families. *Journal of Educational Psychology*, *102*, 43. doi:10.1037/a0016738

Wilson, S. J., & Farran, D. C. (2012, March). *Experimental evaluation of the Tools of the Mind preschool curriculum*. Paper presented at the Society for Research on Educational Effectiveness, Washington, DC.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock–Johnson Tests of Achievement*. Itasca, IL: Riverside.

Supporting Information

Additional supporting information may be found in the online version of this article at the publisher's website:

Table S1. Descriptive Statistics of Analysis Variables

Table S2. SEM Estimates of the Association Between First Grade Mathematics Achievement and Analysis Mediators

Table S3. SEM Estimates of the Association Between First Grade and Age-15 Mathematics Achievement

Table S4. Regression Adjusted Estimates of the Association Between First Grade and Age-15 Mathematics Achievement: Full Models

Table S5. Regression Adjusted Estimates of the Association Between First Grade and Age-15 Mathematics Achievement with EF Measures Averaged Into One Variable

Table S6. Correlations Between Control Variables and Main Analysis Variables

Appendix S1. Information Regarding Study Methods and Measures, and Additional Results From Additional Analyses