Journal of Educational Psychology
2015, Vol. 107, No. 3, 771–788

0022-0663/15/$12.00 http://dx.doi.org/10.1037/edu0000004

Early Childhood Memory and Attention as Predictors of Academic Growth Trajectories

Deborah Stipek and Rachel A. Valentino
Stanford University

Longitudinal data from the children of the National Longitudinal Survey of Youth (NLSY) were used to assess how well measures of short-term and working memory and attention in early childhood predicted longitudinal growth trajectories in mathematics and reading comprehension. Analyses also examined whether changes in memory and attention were more strongly predictive of changes in academic skills in early childhood than in later childhood. All predictors were significantly associated with academic achievement and years of schooling attained, although the latter was at least partially mediated by predictors' effect on academic achievement in adolescence. The relationship of working memory and attention with academic outcomes was also found to be strong and positive in early childhood but nonsignificant or small and negative in later years. The study results provide support for a "fade-out" hypothesis, which suggests that underlying cognitive capacities predict learning in the early elementary grades, but the relationship fades by late elementary school. These findings suggest that whereas efforts to develop attention and memory may improve academic achievement in the early grades, in the later grades interventions that focus directly on subject matter learning are more likely to improve achievement.

Keywords: memory, attention, executive functions, math, reading

Success in school requires many skills. For example, for children to navigate school settings effectively they need to be able to focus their attention on their teacher, complete tasks in the context of many distractions, and inhibit impulsive thinking and behavior. They also need to remember instructions and be able to complete tasks without forgetting critical information.

The importance of these attention and memory skills for academic success is supported by both theory and research. Many theorists have noted that both short-term and working memory is required for the complex cognitive operations involved in learning school subjects such as mathematics and reading, and in the last decade there has been a proliferation of studies demonstrating that several different facets of memory predict academic skills (see Raghubar, Barnes, & Hecht, 2010; Savage, Lavers, & Pillay, 2007). In addition, extant studies have shown significant associations between children's ability to regulate their attention and their academic performance (e.g., Duncan et al., 2007; Kos, Richdale, & Hay, 2006).

The present study examines associations between attention and memory and academic skill development. Specifically, this study assesses how well attention and both short-term memory and working memory in early childhood predict growth trajectories in math and reading comprehension through adolescence and education attainment in young adulthood. Two contrasting models of associations are compared. The first model, which is implicit in the linear correlational analyses in extant research, assumes a continuous association between initial attention and memory at school entry and academic achievement through childhood and adolescence. In this model, growth in attention or memory should predict growth in achievement throughout school. In this case early memory and attention scores would predict academic achievement roughly at the same level through the elementary and middle school grades and beyond. Throughout the article we refer to this as the "continuous" model.

We posit an alternative model in which the direct effects of these underlying cognitive skills on academic growth fade with time in school and that by the upper elementary grades subject-matter skills become more potent predictors of future learning and performance. In this case, early memory and attention would be strong predictors of academic achievement in the lower elementary grades, but the predictions would be weaker or nonexistent by the time students were in the upper elementary and middle school grades. Throughout the article we refer to this as the "fade-out" model. We reason that children's literacy and math skills at school entry predict their later literacy and math skills, but during the first few years of school, children do not have a long history of academic achievement. Thus, their basic cognitive capacities may play a significant role in how well they are able to take advantage of discipline-based instruction. By the upper elementary grades children have well established differences in the discipline-based skills and knowledge that form the foundation for (and presumably affect) future learning. Their academic skills may therefore become a more important factor in how well they develop further discipline-based skills than more generic cognitive capacities.

A second reason to expect the association between attention and memory and academic learning to be stronger in the early grades than in the later grades is that growth in the prefrontal cortex, which is substantially responsible for the development of these

This article was published Online First November 17, 2014.

Deborah Stipek and Rachel A. Valentino, Graduate School of Education, Stanford University.

Correspondence concerning this article should be addressed to Deborah Stipek, Graduate School of Education, Stanford University, Stanford, CA 94305. E-mail: stipek@stanford.edu

basic cognitive functions, is greater during these early years (Thompson & Nelson, 2001). Accordingly, we propose that early memory and attention skills launch children on academic trajectories, which in turn affect future academic performance and attainment. We hypothesized, specifically, that memory and attention around school entry and growth in memory and attention would predict growth in academic performance for children through the early elementary grades, after adjusting for their basic academic skills at school entry, but that any relationship between growth in memory and attention and achievement past the early elementary grades would be weak. We further test a hypothesis that the effect of initial attention and memory on educational attainment in adulthood would be mediated by previous academic performance.

The present study builds on a substantial body of evidence demonstrating associations between academic achievement and short-term and working memory as well as attention. By comparing different cognitive skills and by examining associations in different periods of development the findings of the study could have implications for the nature and timing of interventions. We operationalize these constructs and review this literature below.

Memory and Academic Skills

Short-term memory involves holding a limited amount of information in a very accessible state temporarily (Atkinson & Shiffrin, 1968). There is not complete consistency in definitions of working memory, but most definitions involve the capacity to store, retrieve and manipulate information over short periods of time or while engaging in other cognitively demanding activities. Working memory is generally viewed as the combination of multiple components working together. Baddeley and Hitch (1974) posited that working memory involves a central executive, which directs attention to relevant information, suppresses irrelevant information and inappropriate actions, and coordinates cognitive processes when more than one task must be done at the same time. Although there is substantial unshared variance, working memory is strongly associated with tests of intellectual aptitude (Ackerman, Beier, & Boyle, 2005), presumably because both aptitude and working memory depend on the ability to control attention (Engle, Tuholski, Laughlin, & Conway, 1999). Although there is some dispute about how malleable working memory is (see Klingberg, 2012; Shipstead, Hicks, & Engle, 2012), studies demonstrating effects of working memory training (e.g., Klingberg, 2012; Stepankova et al., 2014) are evidence for its malleability.

A typical strategy for measuring short-term memory in children is to ask them to reproduce a list of numbers or words. Working memory tasks require some manipulation, such as reproducing numbers backward. Within working memory, a distinction has been made between visual-spatial input and the processing of verbal speech input (referred to as the phonological loop; see Baddeley & Hitch, 1974). The present study included only a measure of the phonological loop component of working memory because it operates on verbal information (Ackerman et al., 2005), which is relevant to all academic subjects, and it is the most commonly used measure of working memory in studies examining associations with academic achievement.

Studies linking performance on memory tasks to performance on academic tasks focus mostly on working memory, in some cases using short-term memory as a covariate, but studies have also shown associations between academic skills and short-term memory. The present study therefore casts a wide net by including measures of both short-term and working memory, employing in each case one verbal and one numeric task.

Previous theorists have proposed mechanisms by which both short-term and working memory might affect learning. Gathercole, Brown, and Pickering (2003) suggested that individual differences in the capacity to store and process material in complex tasks can directly affect children's ability to develop knowledge and skills in key domains over the school years. Swanson and Beebe-Frankenberger (2004) proposed, more specifically, that poor working memory capacity compromises one's ability to simultaneously maintain recently retrieved knowledge while integrating it with concurrent external inputs. This may, in turn, limit learning in both reading and mathematics.

Math

Several theorists have argued that working memory is more relevant to mathematical skill development than reading, because mathematical problem solving requires holding information in mind and acting on it to arrive at a solution (Bull & Scerif, 2001; see also Noël, 2009). There are also far more studies that have assessed associations between memory and math skills than between memory and reading skills. Although some researchers consider short-term as well as working memory in their analysis, more attention has been given to working memory as a potential explanation for difficulties in learning math. Blair, Knipe, and Gamson (2008) proposed that working memory might assist with encoding and retrieval of math facts from long-term storage, and Swanson and Jerman (2006) pointed out that the ability to use working memory resources when attempting to reach an answer is important in learning arithmetic. Although the research is not entirely consistent, many previous studies have shown that performance on working memory tasks differentiates children who have math learning difficulties from those who do not (see Raghubar et al., 2010). Swanson and Jerman (2006) found in a meta-analysis that verbal working memory, but not visual-spatial working memory or short-term memory of words or digits, strongly differentiated children diagnosed with math learning disabilities from children with average math abilities.

Studies have also shown significant associations between working memory and math skills in general populations, even after other key variables are covaried. For instance, the relationship holds after variables such as age, short-term memory, reading, and processing speed are held constant (Berg, 2008). The same is true when looking at the relationship conditional on the child's level of fluid intelligence, reading and arithmetic achievement, phonological processing, short-term memory, and inhibition (Swanson & Beebe-Frankenberger, 2004); reading, IQ, perseveration, and inhibition efficiency (Bull & Scerif, 2001); maternal education and child vocabulary (Espy et al., 2004); and reading, age, and IQ (Andersson, 2007).

Reading

Although the emphasis in the literature has been more on math, some extant evidence suggests that both short-term and working memory also differentiate children with reading problems from children who do not have problems learning to read (see meta-analysis by Booth, Boyle, & Kelly, 2010). Swanson and Jerman (2007) found that both short-term and working memory differentiated children with reading difficulties and also that growth in working memory predicted growth in reading skills. The findings on associations between short-term and working memory and reading difficulties are, however, inconsistent (see Booth et al., 2010; Savage et al., 2005), with some studies reporting significant results and others failing to find a significant association.

There is some evidence that reading comprehension, but not reading accuracy, is associated with working memory (Oakhill, Cain, & Bryant, 2003) as well as short-term memory (see Savage et al., 2005). Given the evidence, that memory may be more important for integrating prior knowledge while simultaneously reading and drawing inferences from text than for rote word-reading tasks, the current study focuses on reading comprehension.

Longitudinal Studies

Although most research documenting associations between memory and academic outcomes are based on cross-sectional studies, a handful of longitudinal studies have found that either short-term memory (Bull, Espy, & Wiebe, 2008), working memory (Hitch, Towse, & Hutton, 2001; Monette, Bigras, & Guay, 2011), or both (Hecht, Torgesen, Wagner, & Rashotte, 2001) predict reading or math skills at least a year later. Only one study was found that assessed how well academic achievement predicted later memory. Welsh, Nix, Blair, Bierman, and Nelson (2010) reported that working memory (combined with attention shifting) at the beginning of preschool predicted both literacy and math skills at the end of the year and in kindergarten, with language skills covaried, but early literacy skills did not predict memory and attention, suggesting that memory and attention may have been driving the association. The links were, however, significant in both directions for math.

Summary

From our reading of the literature, there is little support for the view that memory is more strongly associated with math learning than with reading. As mentioned above, although Monette et al. (2011) found a stronger association between working memory and math than reading, many other studies report that both reading and math achievement were predicted by working memory (e.g., Hitch et al., 2001; St. Clair-Thompson & Gathercole, 2006) or both working and short-term memory (Bull et al., 2008). One study of preschool-age children found that working memory predicted early math skills but not preliteracy in their population of American children, but working memory predicted performance in both domains for Chinese children (Lan et al., 2011). And a few studies have found stronger associations between working memory and reading than math (Gathercole, Alloway, Willis, & Adams, 2006; Gathercole et al., 2003; St. Clair-Thompson & Gathercole, 2006). Notwithstanding these studies favoring one academic domain or another, taken as a whole, extant research indicates that performance on working-memory, and to a lesser degree short-term memory tasks, is likely related to learning in both academic domains. Accordingly, the present study includes measures of both short-term and working memory and both math and reading skills.

Attention and Academic Skills

Attention is defined variably and figures in to a number of different literatures, including research on temperament (where it is typically described as an ability to sustain attention; Rothbart & Jones, 1998) and executive functions (where the focus is on attention shifting; Blair & Razza, 2007). Attention, as measured in the present study, is best described as the ability to regulate attention and resist being distracted, as conceptualized in research on attention-deficit/hyperactivity disorder (ADHD). Attention and working memory are not entirely distinct cognitive functions, as working memory involves attention processes (Engle, 2002; Terman, 1916). Engle (2002) explained that working memory capacity reflects an ability to use attention to avoid distraction. Working memory is thus strongly implicated in attention problems associated with ADHD, and there is evidence that working memory training can reduce attention problems in people diagnosed with ADHD (e.g., Beck, Hanson, Puffenberger, Benninger, & Benninger, 2010). Barkley (1997) proposed, moreover, that attention problems and impulsivity can emerge when children are challenged by poor working memory.

A substantial body of clinical literature has shown that children who have difficulties with attention, along with high impulsivity and activity levels, perform relatively poorly in school (see Kos et al., 2006). Indeed, some studies of externalizing problems suggest that attention, which is associated with both aggressive behavior and academic achievement, explains the common finding of significant associations between aggressive behavior and under-achievement (e.g., Barriga et al., 2002: Frick, Kamphaus, Lahey, Loeber, Christ, Hart, & Tannenbaum, 1991; see Hinshaw, 1992).

A few studies have tracked associations between attention and academic achievement longitudinally in nonclinical samples, but similar to research on memory, studies typically do not track academic outcomes beyond a few years. Generally extant research indicates that attention (as conceptualized in the present study) measured in preschool and elementary school predicts academic outcomes at least a year later (Diamantopoulou, Rydell, & Thorell, 2007; Duncan et al., 2007; Martin & Holbrook, 1985; National Institute of Child Health and Human Development Early Child Care Research Network [NICHD ECCRN], 2003; Rudasil, Gallagher, & White, 2010). For example, in an analysis of six longitudinal studies, Duncan et al. (2007) found that attention assessed at school entry significantly predicted reading and math skills in later grades.

Taken together, these studies provide some evidence for a link between attention, as assessed in the current study, and both math and reading skills. Theorists have not proposed comparatively stronger associations between attention and math than reading, as some have proposed for memory.

Education Attainment

To our knowledge, no study to date has examined associations between early memory and attention and children's eventual years of schooling attained. One related study found that "attention span task-persistence" assessed at age 4 years significantly predicted the odds of completing college by age 25, and the effect was not fully mediated by math or reading skills at age 7 or age 21 (McClelland, Acock, Piccinin, Rhea, & Stallings, 2013). The items in their measure of attention span/task persistence, however, focused substantially on children persisting with a task in the face of difficulty (e.g., "persists at a task until successful"). The attention construct assessed in the present study no doubt overlaps with "attention span task-persistence" assessed in the McClelland et al. (2013) study, but their measure also likely taps other variables, such as motivation and self-confidence, whereas our measure is more exclusively tied to attention.

Our interest in education attainment is in part driven by knowledge that eighth grade academic achievement is one of the strongest predictors of college readiness, even more so than skills observed in high school (ACT, 2008). Consistent with our proposed model of diminishing effects of underlying cognitive skills, we hypothesized that any relationship between early memory and attention and years of schooling attained would operate *through* their influence on academic skills developed by early adolescence.

Summary

Extant research provides substantial evidence indicating that both short-term and working memory and attention predict math and reading skills, but much of the existing literature relied on small samples and even the few that were longitudinal typically only tracked children for a few years. Despite the consistency of findings to suggest that there is a relationship between memory and attention and academic gains, extant research imposes a strict linear relationship (a continuous model), without considering the possibility that such a relationship might be piecewise linear—that the relationship might be strong in early years but weaker in later years (a fade-out model).

Further, although a variable closely related to IQ is occasionally included, many previous studies did not account for other potential confounds such as gender, race/ethnicity, and family socioeconomic status (SES). Previous studies are typically of either memory or attention and have not compared the relative predictive strength of different measures of each by including all of them in the same model. And no study to date has compared the effects of memory and attention on academic performance at different stages of children's educational careers or on ultimate years of schooling attained.

The Current Study

The current longitudinal study follows about 6,000 children through the middle school grades to track academic growth trajectories and for a subset of nearly 2,000 students we track years of schooling attained. We also include a piecewise growth analysis to determine whether the link between growth in attention/working memory and math/reading skills is stronger in early than in later childhood. This is done by fitting a spline (i.e., piecewise) hierarchical linear model. We fit one slope between memory/attention and academic skills in early years and another between the two in later years to determine whether the slopes differ between the two age spans. We also conduct an analysis of whether there is a relationship between very early memory and attention skills and ultimate years of schooling attained as an adult, and if so, whether this relationship is mediated by middle school academic skills. All models control for critical covariates that might be correlated with both early memory or attention and academic growth trajectories.

The study addresses four questions: (a) Do initial scores in short-term and working memory (as measured by forward digit span and a verbal memory task and backward digit span, respectively) and in attention predict initial scores and growth in math and reading comprehension? (b) What is the relative strength of these predictors in explaining academic skill development? (c) Do memory and attention predict academic growth trajectories more strongly in the early years of schooling (kindergarten through the third or fourth grade) than in the later years (third or fourth grade through middle school)? (d) Do initial scores in memory and attention predict the years of schooling children ultimately attain as adults, and are such relationships mediated by academic skills in later childhood?

Method

Sample

The current study included the children of the nationally representative sample of female subjects in the original National Longitudinal Survey of Youth (NLSY) panel survey. Analyses were restricted to children who have nonmissing demographic variables and who, at a minimum, have an initial score on the memory, attention, and vocabulary assessments. The method used in this article allows us to use all available outcome data for children, even if there is missing data from one point of data collection. The final sample included 5,873 children, although a few analyses use smaller samples, 4,124 children and 2,416 children for two of the memory measures because they were not administered in all waves of the study. We ran our models on each of these smaller samples even when the full sample was available to ensure that findings were robust across samples. The sample was about equally divided between girls and boys; 29% were Black, 20% Hispanic, 51% were of other ethnic backgrounds. The NLSY study oversampled by race to ensure that the study captured sufficient samples of children from each racial/ethnic group. Inverse probability weighting (probability weights of 1/probability of sample selection) were used in all regression analyses to adjust for oversampling and ensure unbiased parameter estimation. The average highest level of children's mothers' education was 13.5 years, the equivalent of just over 1 year of college, with a standard deviation of 2.5 years. The average age of mothers at the birth of their children was 26.3 years, with a standard deviation of 5.2 years.

Procedures

Unless otherwise specified below, child measures were obtained in direct, one-on-one interactions with a trained experimenter. The ages at which the initial assessments used in the current study were conducted varied–verbal memory and vocabulary when children were between 3 and 5 years old, attention at age 4 or 5, and short-term and working memory at age 7. Children's vocabulary at about age 4, used as a covariate in analyses as a proxy for IQ (Markusic, 2012), was assessed initially at age 3 years and the academic achievement tests were first given at age 5 years.

The NLSY panel survey was designed to collect data and assess individual children on the measures used in this article biennially. Data were collected every other year from 1986 through 2010. Because of the biennial nature of data collection, most children were assessed five or six times between the ages of 5 and 14 years, typically every 2 years. All assessments were conducted in the child's home while mothers completed surveys. To answer the first two questions regarding how initial attention and memory predict academic trajectories, we compute initial scores of our predictors using the first available observation for children as long as it was collected within the first wave that the measure was administered. This means that the initial scores for attention were measured at age 5, for both measures of verbal memory at age 3, and for both measures of digit span at age 7.

Measures

Additional details on the measures and their administration can be found in Chapter 4 of the NLSY Child Handbook (Baker, 1993).

Numeric digit-span memory tasks. A subscale of the Wechsler Intelligence Scales for Children—Revised (WISC–R; Wechsler, 1974) was used to assess memory biennially for children from age seven through 14 years, although in later years of the study children were only assessed through age 11. The measure has been widely used, and there is substantial evidence of its predictive validity (Mott, 1995). There are two parts to the memory for digit span assessment: digits forward, which assesses short-term auditory memory, and digits backward, which assesses working memory. For digits forward children listen to and repeat a sequence of numbers said by the interviewer. Digits backward measures the child's ability to manipulate verbal information while in temporary storage (working memory); children listen to a sequence of numbers and repeat them in reverse order. In both cases the length of each sequence of numbers increases as children correctly respond. Each correct response is worth one point, with a maximum of 14 for each subscore series and 28 for the total score. Internal consistency for the backward digit span task is reported as 0.80 and test–retest reliability as 0.67. For forward digit span internal consistency reliability is reported as 0.83 and test–retest reliability as 0.72 (Williams, Weiss, & Rolfhus, 2003). Where appropriate, this assessment was administered in Spanish. Mean initial scores at age seven were 5.42 ($SD = 1.96$) for digit-span forward and 3.72 ($SD = 1.55$) for digit-span backward. Forward and backward digit span were positively correlated ($r = .35$) but not collinear.

Verbal memory task. Children's memory in response to auditory stimuli was assessed between the ages of 3 and 6 years using the verbal memory task, one of six subscales of the McCarthy Scales of Children's Abilities. The verbal memory task has been shown to correlate with standard IQ tests, such as the Stanford-Binet (e.g., Davis & Rowland, 1974), and analysis of the NLSY Children data revealed that it predicted gender differences related to a verbal test, favoring girls (Mott, 1995). To assess short-term memory children are first asked to repeat words (Part A) or sentences (Part B) said by the interviewer. Working memory is assessed by asking children to listen to and retell the essential aspects of a short story read aloud by the interviewer (Part C). The score children receive for Part A is based on the number of words uttered by the interviewer that the child repeats and for Part B on the number of key words they repeat from the sentence read by the interviewer. Scores on Part A and B are combined into one composite score by NLSY. Only if children reached a minimum combined score for Parts A and B are they administered the story (Part C), in which children are scored on the basis of their ability to recall and articulate key ideas from a story they are read. Reliability for these two measures ranges from 0.84 to 0.89, depending on the measure of reliability used (Roid, 2003). The scores used in the analyses were based on children's performance in comparison to a nationally representative sample. In the current sample the mean initial score for A & B before standardization was 95.53 ($SD = 15.37$), and for Part C it was 97.15 ($SD = 14.29$).

Attention. Items from the Behavior Problems Index (BPI; based on the Achenbach Behavior Problems Checklist; Peterson & Zill, 1986) were used yearly to assess the attention dimension from age 4 through 14 years. The BPI has been found to be a robust predictor of a wide range of family inputs and child behavior problems (Mott, 1995). The hyperactivity subscale, which contains all of the items used in our attention measure, was used as a measure of attention in analyses of NLSY data by Duncan et al. (2007). It was found to predict academic skills through the early elementary grades and has a reliability score of about 0.70 (varies slightly depending on the age of assessment, Zill, 1990). Parents reported on specific behaviors that children might have exhibited in the previous 3 months on a 3-point scale anchored at 1 (*often true*), 2 (*sometimes true*), and 3 (*not true*). Three items from the BPI "hyperactivity" subscale were used ("child has difficulty concentrating," "child is impulsive or acts without thinking," "child is restless, overly active, cannot sit still"). The items were highly correlated to each other (Cronbach's alpha for initial score = .71). A mean of the items was used in analyses, with a higher score reflecting lower attention problems. The mean initial score at age 5 years for the sample used in this article was 2.48 ($SD = 0.50$).

Peabody Picture Vocabulary Test (PPVT). The PPVT was used to assess children's receptive vocabulary for standard American English and to provide an estimate of scholastic aptitude. Test–retest reliability reported in the manual is 0.93 and split-half reliability of internal consistency is 0.94 (Dunn & Dunn, 1981). The scores used in the present analyses were from assessments administered to children at age 5 years. The English language version of the assessment consists of 175 vocabulary items increasing in difficulty. Children listen to a word uttered by the interviewer and then select one of four pictures that best describes the word's meaning. A child's entry point is based on his or her age. A "basal" is established when a child correctly identifies eight consecutive items and a "ceiling" is established when a child incorrectly identifies six of eight consecutive items. A child's raw score is determined by adding the number of correct responses between the basal and ceiling to the basal score. The average initial score for the PPVT before it was standardized was 37.36 ($SD = 18.62$).

Academic achievement. The Mathematics and the Reading Comprehension scales of the Peabody Individual Achievement Test (PIAT; Dunn & Marwardt, 1970) were used to assess academic achievement biennially from age 5 through 14 years. Duncan et al. (2007) reported test–retest reliability for math of 0.74 and for reading comprehension 0.89. The PIAT Mathematics subscale consists of 84 multiple-choice items of increasing difficulty. It begins with early skills such as recognizing numerals and progresses to more advanced concepts in geometry and trigonometry. Children begin the assessment with an age-appropriate item and establish a "basal" score by attaining five consecutive correct responses. A "ceiling" is reached when five of seven items are answered incorrectly. The nonnormalized score used in analyses is equivalent to the ceiling item minus the number of incorrect responses. For the Reading Comprehension subtest children read a sentence and then select one of four pictures that best portrays the meaning of the sentence. Overall there are 66 sentences ordered in a sequence of increasing difficulty. The PIAT tests are optimal for growth modeling because their scores increase steadily as children age and are not standardized or renormed with age (Singer & Willett, 2003). Means and standard deviations for both the math and reading comprehension PIAT scales at each age can be found in Table 2.

Educational attainment. The data on educational attainment were obtained starting when children were age 15 and older. Participants biennially reported on the total number of years of schooling completed until the final year of data collection in 2010. For the purpose of our data analysis, the maximum value reported across the interviews completed was used. This ensures that we account for the possibility that years of schooling would increase from one interview to the next but also allows us to minimize missing values. The sample was restricted to students who were old enough to have completed college, between the ages of 22 and 30 years. Similar to maternal education, years of schooling attained by children is constructed as a continuous variable, such that 12 indicates that the person completed high school, 13 means high school plus 1 year of college, and so on. Consistent with the U.S. average, the mean for the sample was 12.96 ($SD = 2.26$).

Analysis Plan

Growth curve analysis using hierarchical linear modeling (HLM; Raudenbush, Bryk, Cheong, Congdon, & du Toit, 2011) was used to test the three research questions about academic outcomes: (a) Do initial scores in short-term and working memory (verbal and numerical) and in attention predict initial scores and growth in math and reading comprehension? (b) What is the relative strength of these predictors in explaining academic skill development holding the others constant? (c) Do memory and attention predict academic growth more strongly in the early years of schooling than in the later years? We conducted a mediation analysis (MacKinnon, 2008) to answer our final question: (d) Do initial scores in memory and attention predict the years of schooling children attain as adults, and if so, is the relationship mediated by academic skills in later childhood?

Growth curve modeling in HLM was used to answer our first three questions because it allows us to estimate observations over time nested within individuals to describe patterns of academic growth trajectories. More specifically it is a means to model intraindividual differences in intercepts and slopes (Level 1) as well as model these intercepts and slopes as a function of interindividual differences (Level 2). That is, by allowing individuals' intercepts and slopes to vary randomly, we are able to estimate unique growth trajectories for each individual and account for both variation in scores over time within individuals and variation in intercepts and slopes across individuals. Growth curve models are a type of random-coefficient model where time (or in our case age) varies randomly between subjects (Rabe-Hesketh & Skrondal, 2012; Singer & Willett, 2003).

An additional benefit is that it allows for missing outcome data as long as it is ignorably missing across students. When fitting a growth model it is implicitly assumed that each individual's observed record is a random sample of that individual's true growth trajectory (Singer & Willett, 2003). Because a student's current value on a given variable is highly correlated with past and future values of that variable, the risk of this assumption being violated is minimal. HLM adjusts for missing outcome values that occur when a participant missed one or more waves of data collection through weighting and smoothing techniques. As described above, in all HLM models, the nonnormalized math or reading comprehension score was used. Scores were not normalized because if scores are standardized or renormed with age to represent a mean of zero at each time point, it would be impossible to model growth along a constant scale over time (Singer & Willett, 2003). The PIAT tests are optimal for growth modeling because their scores increase steadily as children age. We do, however, standardize all stable predictor variables such as initial attention score, for which we are not trying to model changes over time. We provide interpretations of the coefficients of interest in the findings section below, but as a rule of thumb, results can be interpreted as the number of points gained in the outcome variable (i.e., math or reading comprehension) for every standard deviation increase in the predictor. To ease interpretation, we provide means and standard deviations of math and reading comprehension scores across each age in Table 2.

Continuous hierarchical linear models. The first two research questions, (a) whether initial scores in short-term and working memory and in attention predict initial scores and growth in math and reading comprehension and (b) what the relative strength of these predictors is in explaining academic skill development, were answered using two-level hierarchical linear models (HLM). More specifically, children's individual intercepts and slopes were predicted using time-invariant student characteristics and the initial attention or memory score of interest. Only one predictor of interest (e.g., attention) was entered in the model at a time to answer the first question. To answer the second question these predictors were gradually built into one full model (as presented in Table 5). We include a random effect on both the intercept and age slope to allow each child to have his or her own intercept and growth trajectory.

Our growth curve models using students' initial scores to predict academic trajectories as children age took the following form:

$$\text{Level 1}: Y_{it} = B_{0i} + B_{1i}Age_{it} + r_{it}.$$

$$\begin{aligned}\text{Level 2}: B_{0i} &= \gamma_{00} + \gamma_{01}iEF_i + \left(X_i - \bar{\bar{X}}\right)\delta_0 + \mu_{0i}\\ B_{1i} &= \gamma_{10} + \gamma_{11}iEF_i + \left(X_i - \bar{\bar{X}}\right)\delta_1 + \mu_{1i}\\ r_{ig} &\sim N\left(0, \sigma^2\right); \begin{bmatrix}u_0\\u_1\end{bmatrix} \sim \left[\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}\tau_{00} & \tau_{01}\\ \tau_{10} & \tau_{11}\end{pmatrix}\right].\end{aligned}$$

Y_{it} is the academic achievement outcome variable for child i at time t and is a linear function of the child's age at that time point plus a random error, r_{it}. Because of how the models were estimated, growth rates resemble an annual effect even though data were collected biennially. This is because children entered the data set at different ages, so all ages are represented in the continuous *Age* term as opposed to a more coarse version of the variable only containing every other age. For example, child A may have first been assessed at age 5, while child B may have first been assessed at age 6. Although individual children were only assessed every other year, there are observations available for some students at every age to estimate an average linear growth term across all ages. Through HLM, we are able to estimate an average annual rate of change in our outcome of interest.

In the above model r_{it} can be interpreted as the measurement error associated with the differences between children's true and observed growth trajectories. Children's intercept B_{0i} is their score on the outcome variable of interest at the initial point of observation, while B_{1i} is their slope on the outcome variable meant to represent average growth in achievement as they age (centered at age 5, approximately the first point of observation). Children's intercepts and slopes are explained by their initial predictor variable (e.g., attention score; iEF_i) and a vector of grand-mean centered covariates $(X_i - \overline{\overline{X}})$, including their gender, race/ethnicity, socioeconomic status (SES; as proxied by maternal education), maternal age at birth of child, and initial PPVT score. Further, these intercepts and slopes are allowed to vary randomly for each student through the terms μ_{0i} and μ_{1i}, respectively. Finally, this method relies on the assumption that all errors have a mean of 0 and are normally distributed. τ_{00} and τ_{11} are values of the variances of individual intercepts and slopes, respectively. A deviance test was used to test the null hypothesis that $\tau_{11} = 0$, to determine whether the random effect on the age slope significantly contributed to the model. It did in all models, with $\chi^2(2)$ ranging from 268.90 to 445.84, $p < .0001$, for math and $\chi^2(2)$ ranging from 280.23 to 543.56, $p < .0001$, for reading comprehension. Results of the deviance test suggest that the random intercepts and slopes model specification is preferred over a specification only allowing student intercepts to vary.

Piecewise "fade-out" hierarchical linear models. To test our third question, as to whether changes in attention and working memory (measured by backward digit span) were stronger predictors of changes in academic achievement in earlier years of schooling than later years, we estimated piecewise, or spline regression models using HLM (Rabe-Hesketh & Skrondal, 2012). In these models we use time-varying versions of our predictors of interest (attention and working memory), which were thus entered at the Level 1. We only examined changes in attention and backward digit span (not verbal memory) because only these variables were consistently collected at the same ages that math and reading comprehension were assessed over time.

$$\text{Level 1}: \; Y_{it} = B_{0i} + B_1\, I2 + B_{2i}\, Age1_{it} + B_{3i}\, Age2_{it} + B_4\, Age1xEF_{it} + B_5\, Age2xEF_{it} + r_{it}$$

$$\begin{aligned} \text{Level 2}: \; B_{0i} &= \gamma_{00} + \left(X_i - \overline{\overline{X}}\right)\delta_0 + \mu_{0i} \\ B_{2i} &= \gamma_{20} + \left(X_i - \overline{\overline{X}}\right)\delta_2 + \mu_{2i} \\ B_{3i} &= \gamma_{30} + \left(X_i - \overline{\overline{X}}\right)\delta_3 + \mu_{3i} \end{aligned}$$

Our piecewise HLM procedures allow us to model whether the relationship between attention or memory and math or reading comprehension has a different intercept and slope in early years (age 5 through 9 years) versus later years (age 10 through 14 years). In the above model at the Level 1, B_{0i} represents children's intercept in math or reading comprehension at age 5. The coefficient, B_1, represents children's average intercept in math or reading comprehension at age 10, relative to their intercept at age 5. The $I2$ variable takes the value of zero for ages five through nine, and the value of one after. Coefficients of most interest to us are those on the age variables. The coefficient, B_{2i}, on $Age1$, represents the average annual rate of growth in math or reading comprehension for children from ages five to nine. $Age1$ takes continuous increasing age values (centered at age 5) from zero through four for ages 5 through 9, and then takes a constant value of four for all ages thereafter. The coefficient, B_{3i}, on $Age2$ represents the average annual rate of growth in math or reading comprehension for children from ages 10 through 14. $Age2$ holds values zero for ages 5 through 9, and continuous linear values of one through five for ages 10 through 14. For B_4 and B_5 we interact each of these two age slopes with our predictors of interest (attention and backward digit span). These interactions indicate whether the slopes during the ages captured in $Age1$ and $Age2$ are more or less steep than the average child's slope, as children's attention or working memory increases.

Finally, at Level 2 we let children's individual intercepts at age 5 and both age slopes to vary. We also explain these intercepts with the same set of covariates used in our initial score models. We limit random effects to these three level-two equations and do not include them on our additional time-varying age-by-attention/working memory interactions. We did this to ensure that there was sufficient data to estimate the additional variance components and also because we did not have a strong reason to believe that each Age × Predictor slope would also have sizable variation after accounting for the variation in age slopes across students. Not including random effects on these additional time-varying predictors requires the assumption that the person-specific effect is constant across population members (Singer & Willett, 2003). This assumption is not different from the initial-score models, which automatically impose this assumption because initial scores do not have within-person variation to allow for such a residual at the Level 2.

Years of schooling attained mediation analysis. Finally, to estimate the relationship between attention/each form of memory and children's educational attainment, and to determine whether this relationship is a direct one, or whether it is mediated by children's academic performance in middle school, we used ordinary least squares (OLS) regression to conduct a mediation analysis. Education attainment was constructed to be a linear outcome variable. These models were also estimated using ordered probit specifications, modeling a categorical outcome variable of highest degree attained. Results are robust to both model specifications, so for the sake of parsimony and to provide the most easily interpretable results, the linear results are reported. Further, this outcome provides a more informative sense of not just whether a student graduated by a given age as is seen in previous literature but also years of schooling as a continuous measure of education. These models included the same control variables as the HLM models (initial PPVT, sex, race, SES, and maternal age at birth of child) as well as a vector of indicator variables of the year the child was born to account for the fact that students born in earlier years might attain less schooling than students born in later years due to exogenous factors such as increased access to higher education or shifts in the economy. These model results thus reflect within year-of-birth effects, such that students who were born in the same year are compared to each other to derive estimates of interest. Robust standard errors were used. To conduct our mediation analysis, we followed procedures outlined by MacKinnon (2008).

Figure 1 outlines a single mediator model. Here, three regression equations are estimated to assess mediation as illustrated in the figure.[1]

$$(a) Y = B_1 + cX + e_1$$

$$(b) Y = B_2 + c'X + bM + e_2$$

$$(c) M = B_3 + aX + e_3$$

Where Y is the dependent variable (years of schooling attained), X is the predictor (initial attention or initial memory), and M is the mediating variable (middle school academic scores). Further, c represents the total effect between our predictor of interest and years of schooling attained, b represents the relationship between the mediator (middle school academic score) and years of schooling attained, adjusted for the effects of the predictor (attention or memory), a is a parameter that captures the relationship between the predictor and the mediator, and finally c' captures the direct effect that represents the relationship between the predictor and years of schooling after accounting for the mediator (MacKinnon, 2008). This final parameter, c' is our main coefficient of interest, as it indicates what the relationship between initial attention or memory and years of schooling attained is after accounting for middle school academic score. We calculate indirect (or mediated) effects for all models as well ($c - c'$ or $\hat{a}\hat{b}$). The coefficient for the indirect effect represents the change in Y for every standard deviation change in the predictor (e.g., attention) that is mediated by one's middle school academic score. We calculate the standard errors for our indirect effects using the following formula: $se = \sqrt{\hat{a}^2 s_{\hat{b}}^2 + \hat{b}^2 s_{\hat{a}}^2}$ (MacKinnon, 2008). We use children's middle school math score as our mediator. We also ran models using middle school reading comprehension score as our mediator. Results are robust to both specifications, likely because these two variables are highly correlated, and on average both capture "middle school achievement." All variables (initial attention/memory, middle school math score, and years of schooling attained) are stable within students.

Results

Descriptive Statistics

Means and standard deviations for all variables were included in our measures section above. Table 1 displays the correlations among all stable predictor variables and math and reading comprehension scores at each grade. The correlation between scores on the two verbal memory tasks (A & B: Words, and C: Story) were the most highly correlated (r = .59), followed by backward and forward digit span (r = .35). The remaining variables were generally correlated at about 0.20 or slightly less. No variables were correlated enough to be concerned about multicollinearity in our models. A few negative correlations were predictable from previous research; for example black and Hispanic children performed lower on many of these measures than did whites. Further, being female was positively correlated with scores on all attention and memory measures, as well as initial PPVT scores.

Table 2 displays means, standard deviations, minimum and maximum values, and skew of the time-varying variables used in our piecewise linear growth models across each age. Forward digit span was included as a point of reference for backward digit span, although it is not covaried in our piecewise models. As can be seen, these scores increase as children age. Most of the variables have symmetric distributions (as evidenced by skew values between −0.5 and 0.5), but in a handful of grades some of the measures are moderately skewed (as reflected by skew values between −1 and −0.5 or 0.5 and 1). A visual inspection of histograms indicates no major concerns. Table 2 will aid in the interpretation of coefficients in the results section, particularly by looking at the standard deviations of the outcome variables. All predictors were standardized prior to entry into our models, but as indicated earlier, the outcome variables were not. All coefficients can therefore be interpreted as the unit change in the dependent variable for every standard deviation change in the predictor variable. These descriptive tables reflect prestandardized values for all variables.

Continuous Growth Models

Research Question 1: Individual predictors. Tables 3 and 4 present findings from analyses answering the first research question about whether initial memory and attention at school entry predict initial scores and growth in math and reading comprehension over time, respectively. Each column of the tables represents a separate model using a different initial score to predict children's initial achievement outcome and average rate of change as they age. The degree to which these variables predict intercepts and slopes is indicated by initial score and Initial Score × Age, respectively. The initial digit span model (Column 2) includes both initial forward digit span and initial backward digit span. Both

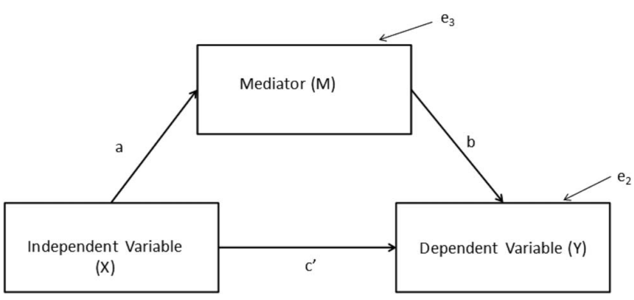

Figure 1. Mediation model used to answer Question 4.

[1] Because we run four separate mediation models, the Bonferroni method would suggest using a critical value of p = .0125 to determine significance. It is the most conservative method that can be used to control for familywise error rates (the probability of making a type I error when performing multiple hypothesis tests). However, the Bonferroni correction introduces a tradeoff by increasing the likelihood of making a type II error and is concerned with the hypothesis that all null hypotheses are true simultaneously (which is not our main question of interest). Further if we run the path c' model (which estimates the effect of the predictor of interest on years of schooling attained after adjusting for middle school scores) with all four memory and attention predictors in the same model, all four measures maintain their level of significance in predicting years of schooling attained (most notably, attention remains significant at the p = .05 level). We therefore do not believe there is a threat of a type I error across these four models. For these reasons we maintain standard critical values to determine statistical significance.

Table 1
Correlation Matrix

Variable	1	2	3	4	5	6	7	8	9	10	11
1. Initial Attention	1										
2. Initial Digit Span (Forward)	.095*	1									
3. Initial Digit Span (Backward)	.110*	.350*	1								
4. Initial Verbal Memory (C–Story)	.066*	.116*	.134*	1							
5. Initial Verbal Memory (A&B–Words)	.087*	.262*	.194*	.585*	1						
6. Initial PPVT	.117*	.184*	.200*	.191*	.200*	1					
7. Female	.113*	.041*	.058*	.088*	.144*	.018*	1				
8. Black	−.067*	.037*	−.094*	−.013	.036*	−.274*	.006	1			
9. Hispanic	−.009*	−.111*	−.048*	−.064*	−.132*	−.147*	−.005	−.297*	1		
10. Mother's Highest Degree Attained	.132*	.159*	.148*	.148*	.157*	.273*	−.011*	.023*	−.152*	1	
11. Mother's Age at Birth of First Child	.159*	.028*	.057*	.034*	.000	.182*	−.005	−.095*	.015*	.299*	1
Math Age 5	.119*	—	—	.264*	.248*	.312*	.053*	−.167*	−.131*	.245*	.155*
Math Age 6	.112*	—	—	.171*	.185*	.331*	.023	−.189*	−.091*	.270*	.167*
Math Age 7	.152*	.266*	.290*	.239*	.296*	.372*	.017	−.187*	−.114*	.320*	.189*
Math Age 8	.150*	.329*	.428*	.206*	.248*	.380*	−.003	−.271*	−.079*	.273*	.165*
Math Age 9	.185*	.270*	.339*	.187*	.281*	.365*	.001	−.226*	−.088*	.302*	.251*
Math Age 10	.166*	.293*	.403*	.196*	.222*	.378*	−.027	−.256*	−.072*	.295*	.224*
Math Age 11	.171*	.274*	.335*	.223*	.272*	.378*	−.032	−.241*	−.112*	.339*	.270*
Math Age 12	.144*	.275*	.366*	.169*	.200*	.359*	−.047*	−.262*	−.116*	.315*	.255*
Math Age 13	.156*	.250*	.342*	.201*	.228*	.391*	−.040*	−.269*	−.094*	.303*	.237*
Math Age 14	.168*	.257*	.347*	.258*	.229*	.375*	−.046*	−.277*	−.099*	.316*	.257*
Reading Comprehension Age 5	.119*	—	—	.260*	.305*	.310*	.095*	−.018	−.155*	.342*	.166*
Reading Comprehension Age 6	.121*	—	—	.160*	.209*	.283*	.102*	−.034*	−.086*	.267*	.176*
Reading Comprehension Age 7	.135*	.259*	.243*	.167*	.276*	.279*	.093*	−.107*	−.075*	.266*	.138*
Reading Comprehension Age 8	.142*	.339*	.374*	.157*	.232*	.354*	.090*	−.179*	−.066*	.236*	.107*
Reading Comprehension Age 9	.156*	.269*	.308*	.203*	.285*	.355*	.074*	−.216*	−.083*	.273*	.128*
Reading Comprehension Age 10	.161*	.294*	.346*	.186*	.229*	.392*	.047*	−.218*	−.062*	.294*	.102*
Reading Comprehension Age 11	.153*	.270*	.317*	.227*	.311*	.389*	.023	−.238*	−.064*	.321*	.118*
Reading Comprehension Age 12	.157*	.297*	.344*	.255*	.230*	.415*	.042*	−.255*	−.054*	.291*	.125*
Reading Comprehension Age 13	.178*	.258*	.307*	.207*	.269*	.417*	.030	−.279*	−.056*	.305*	.129*
Reading Comprehension Age 14	.174*	.270*	.320*	.284*	.253*	.405*	.022	−.292*	−.042*	.313*	.141*
M	2.48	5.42	3.72	97.15	95.53	37.36	0.497	0.292	0.202	12.96	26.33
SD	0.50	1.96	1.55	14.29	15.37	18.62	0.500	0.460	0.400	2.26	5.21

Note. Dashes indicate that data were not collected. PPVT = Peabody Picture Vocabulary Test (Dunn & Dunn, 1981).
* $p < .05$.

were also included to assess rate of change with age. This was done to determine whether working memory (backward digit span), which has been found to be the stronger predictor of academic skill development, is robust to the inclusion of a measure of short-term memory (forward digit span). Our findings show that each of the predictor variables measured initially (attention, digit span forward and backward, and both measures of verbal memory) significantly predicted children's initial scores in both math and reading comprehension. The one exception is Part C of the initial verbal memory subscale, where children retell the essential aspects of the short story. This measure significantly predicted initial math performance but not reading comprehension after adjusting for controls. All but one of the predictor variables also significantly predicted average growth in both math and reading comprehension as children aged. The exception was forward digit span, which did not significantly explain growth in math.

Interpreting the practical significance of these coefficients can be difficult because standardized outcome variables were not used for the purpose of growth modeling. It is possible, however to get a sense of how big of an effect these coefficients reflect using inferences from Table 2, with knowledge that on average the standard deviation for math scores is nine points and a standard deviation for reading comprehension is 10 points. For the sake of parsimony, we focus on interpreting rates of growth here. Children who were one standard deviation higher in initial attention gained anywhere from 0.01 to 0.02 more standard deviations per year in math. Thus, a one-standard-deviation increase in initial attention predicted up to about a fifth of a standard deviation ($9 \times .02$) in additional average math score growth over the 9-year time span. The findings are similar for initial backward digit span after adjusting for forward digit span. For every one standard deviation children scored higher on backward digit span, children gained approximately 0.14 more standard deviations in math from ages 7 through 14. For initial verbal memory scoring one standard deviation higher on Parts A&B of the verbal memory measure resulted in just over 0.13 standard deviations of additional growth between ages 5 and 14 compared to the average student in the sample. Similarly, scoring one stand deviation higher in Part C corresponded to growing about 0.10 more standard deviations, cumulatively, through age 14.

The findings for reading comprehension did not deviate much from the findings for math. Children who initially scored one standard deviation higher in attention grew by approximately 0.01 standard deviations more per grade in reading comprehension than the average, or a tenth of a standard deviation more cumulatively from ages 5 through 14. The point estimate on backward digit span

Table 2
Descriptive Statistics of Time-Varying Variables, by Age

Variable	Age (years) 5	6	7	8	9	10	11	12	13	14
Math										
M	15.079	18.59	25.085	32.672	39.92	45.283	49.328	52.263	54.855	56.22
SD	4.714	6.261	8.513	9.758	10.614	9.652	10.429	10.402	11.009	11.590
Min	7	3	1	1	0	0	6	10	0	0
Max	30	43	74	66	84	84	84	84	84	84
Skew	1.084	1.018	0.617	0.119	−0.253	−0.386	−0.215	−0.118	−0.349	−0.154
Reading Comprehension										
M	18.861	20.258	25.864	33.076	38.168	42.806	46.354	49.485	51.652	54.092
SD	3.319	4.512	7.838	9.534	10.066	10.222	10.886	11.029	11.875	12.018
Min	18	18	0	12	0	0	15	14	0	0
Max	47	46	69	68	78	84	84	82	84	84
Skew	0.018	−0.313	−0.024	0.437	0.005	−0.078	−0.025	−0.092	−0.171	−0.24
Attention										
M	2.506	2.503	2.47	2.472	2.487	2.499	2.529	2.53	2.551	2.545
SD	0.640	0.498	0.541	0.523	0.553	0.531	0.513	0.512	0.547	0.537
Min	1	1	1	1	1	1	0.333	0.333	0.667	1
Max	3	3	3	3	3	3	3	3	3	3
Skew	−0.945	−0.856	−0.723	−0.831	−0.825	−0.901	−0.932	−0.994	−1.094	−1.11
Backward Digit Span										
M	—	—	3.35	3.824	4.36	4.779	5.134	5.434	5.48	6.014
SD	—	—	1.365	1.486	1.588	1.692	1.876	1.955	2.062	2.113
Min	—	—	0	0	0	0	0	0	1	2
Max	—	—	9	14	12	12	13	13	12	13
Skew			−0.05	0.489	0.637	0.543	0.654	0.717	0.732	0.601
Forward Digit Span										
M	—	—	5.184	5.475	5.934	6.38	6.79	7.12	7.081	7.606
SD	—	—	1.833	1.943	1.999	2.114	2.174	2.194	2.013	2.308
Min	—	—	0	0	0	0	0	0	2	2
Max	—	—	13	14	14	14	14	14	14	14
Skew			0.603	0.609	0.476	0.352	0.355	0.153	0.195	0.262
N	2,882	2,991	2,867	3,275	3,481	3,416	3,426	3,244	3,163	2,949

Note. The table includes variables that were collected at each wave of study for all or some portion of the sample. Means and standard deviations are prestandardization (attention and digit span were standardized for the purpose of our hierarchical linear modeling analyses). Although forward digit span was not used in the piecewise growth analyses, it is included in this table to provide a reference for backward digit span. When skew is between −0.5 and 0.5 the distribution is approximately symmetric. When it is between −1 and −0.5 or 0.5 and 1 it is moderately skewed. Dashes indicate that data were not collected.

indicates that for every one standard deviation increase in students' scores, they grew a total of 0.07 standard deviations more than average in reading comprehension from 7 through 14 years old. Finally, based on rough calculations, a one standard deviation increase in initial verbal memory (on either measure) predicted about a 0.17 standard deviation increase in reading comprehension from ages 5 through 14.

Research Question 2: Inclusion of multiple predictors. The next analysis set out to answer the question about the relative strength of these predictors in explaining academic skill development when multiple predictors were included in the same model. Table 5 summarizes findings from a set of models that iteratively add additional predictors to the model. Predictors are entered at level two to predict children's intercepts and rates of growth in math and reading skills as they age. Models 1a and 1b include both initial attention and initial forward and backward digit span. Model 2 includes both measures of verbal memory, while Model 3 includes all of five measures in the same model. Because the inclusion of initial verbal memory (A&B) leads to a much smaller sample size, results from Model 1a are not directly comparable to results presented in Models 2 and 3. For this reason we included Model 1b, which is the same specification as 1a but restricts the sample to that of Models 2 and 3. Given that the correlations between these predictors are generally low, and because the coefficients of these multiple-predictor models do not change substantially from the models that include each of these variables individually, we are not concerned about multicolinearity.

Model 1 shows that the findings are robust to the inclusion of both forms of initial digit span and attention. The point estimates largely do not change for predicting rates of growth in both math and reading comprehension from the models in which they are included alone. All coefficients on growth maintain their significance and generally maintain their magnitude from Models 1b and 2 to Model 3. In Model 2, verbal memory Part A&B but not Part C remains a significant predictor of both children's intercepts and rates of growth in math. For reading comprehension, the reverse is generally true. Finally, Model 3 includes all initial score predictors. These models demonstrate that initial backward digit span and attention remain strong predictors of growth in both math and reading comprehension, even in the models of smaller sample sizes. Further, in terms of initial verbal memory, Part C remains a significant predictor of growth in reading comprehension, which

Table 3
HLM Models Using Initial Scores to Predict Students' Intercepts and Slopes in Math

Variable	Attention b (SE)	Digit Span Forward and Backward b (SE)	Verbal Memory (A and B) b (SE)	Verbal Memory (C) b (SE)
Initial Scores of Predictor Variables				
Intercept	14.842*** (0.100)	15.981*** (0.110)	16.937*** (0.178)	16.556*** (0.223)
Initial Math Scores				
Initial Score	0.376*** (0.101)	1.349*** (0.113)	0.940*** (0.154)	0.619*** (0.186)
Initial Backward Digit Span		1.722*** (0.123)		
Growth in Math Scores				
Age	5.438*** (0.020)	5.487*** (0.022)	4.777*** (0.032)	4.691*** (0.041)
Initial Score × Age	0.088*** (0.020)	0.028 (0.023)	0.127*** (0.028)	0.098** (0.034)
Initial Backward Digit Span × Age		0.161*** (0.025)		
N (Children)	5,873	5,873	4,124	2,416
N (Observations)	24,117	24,117	14,168	8,398
Tau[1,1] (intercept)	9.471	6.278	8.478	8.658
Tau[2,2] (slope)	0.431	0.441	0.352	0.377

Note. HLM = hierarchical linear modeling. In the digit span forward and backward models, the first initial score presented is for forward digit span, while the second refers to backward digit span. All models include controls (Hispanic, Black, maternal highest degree, sex, maternal age at child's birth, and initial Peabody Picture Vocabulary Test score). Controls are grand-mean centered and used to explain child intercepts and age slopes. Age is centered at 5 years of age, since this is the first age at which children are assessed in math and reading comprehension.
** $p < .01$. *** $p < .001$.

seems intuitive given that an early ability to recall key concepts from a story is in many ways an emerging skill in reading comprehension.

Fade-Out Growth Models

Research Question 3. Table 6 presents findings from our piecewise HLM growth models, which were used to answer our research question as to whether changes in attention and working memory more strongly predict academic growth trajectories in early years of schooling (Age 1: ages 5 through 9 years) than later years (Age 2: ages 10 through 14). From the coefficients on Age1 and Age2, we see that growth in math and reading comprehension is, on average, faster from ages 5–9 than it is from ages 10–14, but still significant and positive during both time points. From the

Table 4
HLM Models Using Initial Scores to Predict Students' Intercepts and Slopes in Reading Comprehension

Variable	Attention b (SE)	Digit Span Forward and Backward b (SE)	Verbal Memory (A and B) b (SE)	Verbal Memory (C) b (SE)
Initial Scores of Predictor Variables				
Intercept	15.905*** (0.098)	16.963*** (0.109)	18.558*** (0.182)	18.657*** (0.232)
Initial Reading Comprehension Scores				
Initial Score	0.250* (0.100)	1.364*** (0.112)	0.795*** (0.157)	0.163 (0.194)
Initial Backward Digit Span		1.473*** (0.122)		
Growth in Reading Comprehension Scores				
Age	4.886*** (0.020)	4.928*** (0.023)	4.238*** (0.033)	4.132*** (0.042)
Initial Score × Age	0.107*** (0.020)	0.072** (0.024)	0.185*** (0.029)	0.196*** (0.036)
Initial Backward Digit Span × Age		0.107*** (0.026)		
N (Children)	5,873	5,873	4,124	2,416
N (Observations)	24,117	24,117	14,168	8,398
Tau[1,1] (intercept)	6.855	4.29	7.576	9.883
Tau[2,2] (slope)	0.551	0.575	0.46	0.427

Note. HLM = hierarchical linear modeling. In the digit span forward and backward models, the first initial score presented is for forward digit span while the second refers to backward digit span. All models include controls (Hispanic, Black, maternal highest degree, sex, maternal age at child's birth, and initial Peabody Picture Vocabulary Test score). Controls are grand-mean centered and used to explain child intercepts and age slopes. Age is centered at 5 years of age, since this is the first age at which children are assessed in math and reading comprehension.
* $p < .05$. ** $p < .01$. *** $p < .001$.

Table 5
Models Using Initial Scores to Predict Students' Intercepts and Slopes in Math and Reading Comprehension

	Math				Reading Comprehension			
	Model 1a	Model 1b	Model 2	Model 3	Model 1a	Model 1b	Model 2	Model 3
Variable	*b* (*SE*)	*b* (*SE*)	*b* (*SE*)	*b* (*SE*)	*b* (*SE*)	*b* (*SE*)	*b* (*SE*)	*b* (*SE*)
Initial Scores								
Intercept	16.002*** (0.111)	17.804*** (0.254)	16.499*** (0.224)	17.653*** (0.259)	16.973*** (0.109)	19.830*** (0.262)	18.564*** (0.232)	19.718*** (0.269)
Initial Digit Span Forward	1.339*** (0.113)	1.041*** (0.222)		0.950*** (0.227)	1.360*** (0.112)	1.406*** (0.230)		1.265*** (0.235)
Initial Digit Span Backward	1.711*** (0.124)	1.767*** (0.251)		1.714*** (0.252)	1.467*** (0.122)	1.377*** (0.260)		1.362*** (0.261)
Initial Attention	0.232* (0.098)	0.123 (0.169)		0.103 (0.169)	0.113 (0.097)	0.147 (0.175)		0.128 (0.175)
Initial Verbal Memory (A and B–Words)			0.800*** (0.236)	0.463 (0.250)			1.245*** (0.244)	0.778** (0.259)
Initial Verbal Memory (C–Story)			0.214 (0.221)	0.233 (0.226)			−0.474* (0.229)	−0.321 (0.234)
Growth in Scores								
Age	5.495*** (0.022)	4.834*** (0.046)	4.685*** (0.041)	4.811*** (0.047)	4.937*** (0.023)	4.272*** (0.048)	4.128*** (0.042)	4.238*** (0.049)
Initial Digit Span Forward × Age	0.025 (0.023)	0.089* (0.041)		0.072 (0.041)	0.069** (0.024)	0.048 (0.042)		0.035 (0.043)
Initial Digit Span Backward × Age	0.157*** (0.025)	0.146** (0.046)		0.138** (0.046)	0.102*** (0.026)	0.186*** (0.048)		0.170*** (0.048)
Initial Attention × Age	0.081*** (0.020)	0.102*** (0.030)		0.096** (0.030)	0.100*** (0.020)	0.090** (0.031)		0.083** (0.031)
Initial Verbal Memory (A and B) × Age			0.129** (0.043)	0.077 (0.046)			0.082 (0.045)	0.041 (0.047)
Initial Verbal Memory (C) × Age			0.027 (0.041)	0.045 (0.042)			0.150*** (0.043)	0.140** (0.043)
N (Children)	5,873	2,416	2,416	2,416	5,873	2,416	2,416	2,416
N (Observations)	24,117	8,398	8,398	8,398	24,117	8,398	8,398	8,398
Tau[1,1] (intercept)	6.238	7.947	8.358	7.703	4.295	8.083	9.133	7.942
Tau[2,2] (slope)	0.434	0.303	0.368	0.291	0.566	0.382	0.427	0.358

Note. All models include controls (Hispanic, Black, maternal highest degree, sex, maternal age at child's birth, and initial Peabody Picture Vocabulary Test score). Controls are grand-mean centered and used to explain child intercepts and age slopes. Age is centered at 5 years of age, since this is the first age at which children are assessed in math and reading comprehension.
* $p < .05$. ** $p < .01$. *** $p < .001$.

Table 6
Piecewise Regression Models: Attention and Backwards Digit Span as Predictors of Math and Reading Comprehension

	Math		Reading Comprehension	
	Attention	Digit Span	Attention	Digit Span
Variable	*b* (*SE*)	*b* (*SE*)	*b* (*SE*)	*b* (*SE*)
Intercept 1 (age 5)	8.619*** (0.347)	11.597*** (0.429)	12.586*** (0.364)	15.108*** (0.450)
Intercept 2 (bump at age 10)	3.014*** (0.162)	1.883*** (0.230)	1.810*** (0.177)	1.152*** (0.243)
Age 1 (age 5–9)	7.774*** (0.105)	7.037*** (0.132)	6.459*** (0.112)	5.822*** (0.138)
Age 2 (age 10–14)	2.964*** (0.047)	3.150*** (0.100)	2.883*** (0.051)	2.929*** (0.107)
Age 1 × Attention or Digit Span	0.110*** (0.024)	0.723*** (0.030)	0.163*** (0.026)	0.574*** (0.031)
Age 2 × Attention or Digit Span	0.009 (0.034)	−0.384*** (0.054)	0.044 (0.037)	−0.352*** (0.058)
N	5,873	5,873	5,873	5,873
T11	16.175	16.175	25.71	25.71
T22	0.799	0.799	1.831	1.831

Note. Controls (Hispanic, Black, maternal highest degree, sex, maternal age at child's birth and initial Peabody Picture Vocabulary Test score) are grand-mean centered and used to explain child intercepts and age slopes.
*** $p < .001$.

coefficients on both Age × Attention and Age × Digit Span interactions, there is a clear pattern. Changes with age in both attention and backward digit span are positively related to changes in math and reading comprehension as children advance through the elementary grades, such that higher scores on both of these measures in a given grade predict steeper gains from grade to grade in math and reading comprehension than students would have realized had they performed lower in attention or working memory in that grade. However, we see a different pattern in the later grades. Here, changes in attention do not predict changes in academic skill development above and beyond the average rate of growth (captured by the coefficient Age 2). One possible explanation for this finding is that children grow a great deal in attention in early years, but then slow or flatten in growth in later grades. Interestingly, the relationship between changes in working memory and both math and reading comprehension in later grades is negative, indicating that increases in working memory capacity in later childhood is actually related to less steep rates of increase in academic outcomes as children age. The size of the coefficient is less than half of the size of the coefficient in early childhood, so is not large enough to cancel out the initial effect. The age at which we split the slopes was initially designed to equalize the number of years on either side. We conducted the same analyses splitting the slopes on both sides of age 10 (ages 9 and 11) to test the robustness of our results; results were consistent across all age breaks used to define earlier and later childhood.

Mediation Analysis

Research Question 4: Years of schooling attained. Figure 2 presents results from our mediation analysis used to answer our final question as to whether initial scores in memory and attention predict years of school children attain as adults and whether this relationship is mediated by academic skills in later childhood. First, we note that initial attention ($B = 0.189$, $p < .01$), backward digit span ($B = 0.308$, $p < .001$), and verbal memory A&B (words; $B = 0.213$, $p < .001$) and C (story; $B = 0.203$, $p < .01$) were significant independent predictors of years of schooling attained. These findings are captured in the total effect. We then followed steps outlined by MacKinnon (2008) to determine whether it was *through* middle school academic ability that these variables had their influence on years of schooling attained. We calculated direct effects, or the portion of the total effect that is due to our predictors of interest, after adjusting for the mediator variable, and indirect effects, or the portion of the total effect that is due to the mediator.

The results show that with the exception of attention, all of the direct effects are nonsignificant after including middle school academic score as a mediator. The point estimates also shrink sizably, such that the indirect effect (e.g., working memory's effect on years of schooling attained through middle school academic performance) captures anywhere from 61%–70% of the total effect. Although the significance of these variables disappears, the coefficients do not shrink to zero, so it is difficult to determine with certainty whether the finding reflects full or partial mediation. Attention, on the other hand, maintains significance, and the direct effect did not shrink as much from the total effect as it did in the memory models. Findings from the attention model indicate that about 60% of the effect of attention on years of schooling is a direct one, while about 40% of its effect is mediated by middle school academic performance. Both the direct and indirect effects are significant, suggesting partial mediation. The retained significant direct effect of attention (albeit smaller than the total effect) but not of any of our measures of memory provides some evidence that our measures of memory (mainly working memory) have their impact on years of schooling attained almost entirely through their impact on the development of academic skills, while attention may remain a key ingredient to behaviors such as persistence in school.

Discussion

The present study makes several contributions to the literature. First, it provides evidence on associations between academic achievement and attention and memory assessed on an array of tasks, with a large sample of children, over an extended period of time, and with extensive covariates. In addition to assessing how well these basic cognitive skills at school entry predicted growth in academic learning independently, the study compared the strength of the independent contributions of numerical and verbal short-term and working memory and attention.

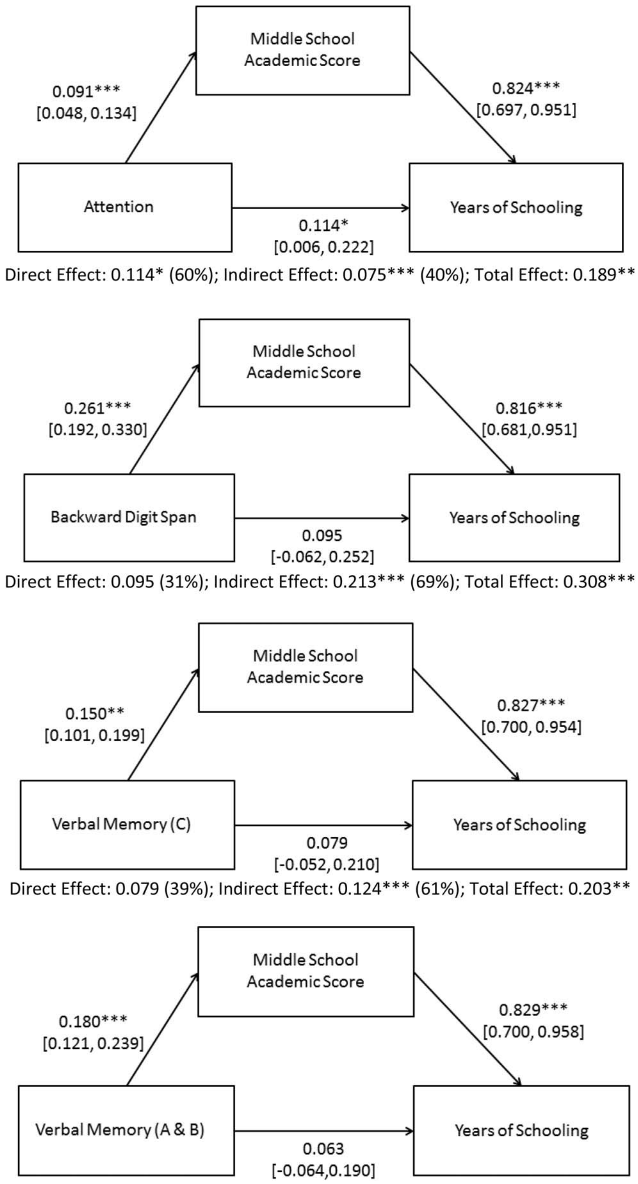

Figure 2. Middle school academic score takes the value of their math score. In each of the four mediation models, initial attention or memory (X) and middle school academic score (M) are standardized, but years of schooling attained is not standardized. For this reason, the coefficient on path a is in standard deviation units, but the coefficients on paths b and c' represent the unit change in years of schooling attained for every standard deviation increase in X or M, respectively. * $p < .05$. ** $p < .01$. *** $p < .001$.

The study posited two competing models for the link between memory and attention and academic achievement—one in which the underlying cognitive capacities exert influence on learning continuously through at least middle school and one in which the influence is strong in the early years of school but fades by late elementary school. Evidence related to these two models is important to examine because they have implications for the nature and timing of intervention, as is discussed below.

Predicting Achievement Gains With Memory and Attention

With regard to our first question, all but one of the memory and attention variables assessed in early childhood were strong predictors of both math and reading comprehension skills at the same age, and they predicted the trajectory of children's academic skill development through the elementary grades, over and above the effects of our control variables. The findings were remarkably similar for math and reading and were generally consistent with studies mentioned above (e.g., Bull et al., 2008; Diamantopoulou, Rydell, & Thorell, 2007; Duncan et al., 2007; Hitch et al., 2001; Martin & Holbrook, 1985; NICHD ECCRN, 2003; Rudasil et al., 2010; St. Clair-Thompson & Gathercole, 2006). The findings are all the more impressive for comprehension given that a test of vocabulary (PPVT), which is strongly implicated in comprehension, was held constant. Taken together, this study and existing research strongly suggests that these basic cognitive capacities are relevant to both subject areas and are likely to be relevant to learning in other subjects, such as history and science, that employ reading and math.

Comparison of Predictors

Understanding the relative importance of different basic cognitive skills can be useful in designing interventions that target the most salient early predictors of academic development. The results of analyses addressing our second question indicated that attention and working memory were the most robust predictors of academic growth trajectories for both math and reading comprehension. Neither the numerical (digit forward) nor the verbal (word recall) short-term memory tasks predicted math or reading skills when the other cognitive skills were included in the analysis.

The study, however, revealed some notable differences in the two working memory tasks. The backward digit span task predicted both math and reading comprehension trajectories over and above the other cognitive measures, but the working memory task requiring children to recall details of a story predicted growth trajectories in reading comprehension but not in math. Given that reading comprehension involves recalling the main ideas of text, the predictive strength of the story recall task is not surprising. The findings nevertheless suggest the importance of attending to the particular qualities of the cognitive task being used as well as the specific aspects of academic achievement being predicted. On the basis of these findings, the backward digit-span task appears to measure a robust cognitive skill that has broad effects on learning, whereas the story recall task may measure skills that have implications for a more limited set of academic skills.

Continuous Versus Fade-Out Models

With regard to the third question we found, as predicted, that changes in attention and working memory predicted increases in academic skills as children advanced through the early elementary grades but not through the late elementary and middle-school grades. The difference was robust to several analytic strategies and could not be explained by differences between these two age spans in the amount of growth in attention or working memory. The results suggest that these basic cognitive skills may play a greater role in the development of math and reading comprehension skills for young elementary-grade children than for older children. The more frequent inclusion of young children in extant studies may explain why they consistently find significant associations between attention and working memory and academic trajectories. Our findings suggest that these findings should not be generalized to older populations.

One possible explanation for the waning effect of the basic cognitive skills is that as children advance in school they develop stable and entrenched individual differences in their academic skills and that, because subject-matter learning builds on previous learning, it is increasingly affected by extant skill levels. The importance of subject-matter knowledge may intensify in middle school, when children typically have different teachers and classes for different subject matter. Accumulated differences in performance histories are also associated with differences in self-confidence and motivation, which play a role in future learning. Children who have a history of low achievement tend to develop negative perceptions of their ability and may as a result exert less effort (see Stipek, 2002). There is ample evidence that with age, children increasingly attend to how their performance compares to peers and incorporate normative information in their judgments of their competencies (Butler & Ruzany, 1993; Ruble, Grosovsky, Frey, & Cohen, 1992). Perhaps because they are paying more attention, their competence beliefs and academic achievement values become more stable as they advance through elementary school and their beliefs in their own academic competence become highly associated with their actual performance (Wigfield et al., 1997). In summary, we speculate that for older children, current subject-matter skills and accompanying motivation-related beliefs may overwhelm the effects of underlying cognitive skills, such as attention and working memory, on how well new subject-matter material is learned.

One finding that was not predicted was that increases in working memory capacity in later childhood were related to less steep rates of increase in academic outcomes as children aged, although the size of the coefficient is less than half of the size of the coefficient in early childhood, so not large enough to cancel out the initial effect. One possible explanation for the negative association is that most children were close to the ceiling of backward digit span by age 11. (The gain in the number of digits children could repeat backward increased from 3.35 to 4.78. between ages 7–10 but only from 5.13 to 6.01 between ages 11–14.) The children whose digit spans continued to increase in later childhood had lower scores to begin with (i.e., their development of backward digit span was delayed); they also had lower academic skills and presumably slower academic growth. Gains in working memory would therefore be associated with slower growth in academic skills.

Education Attainment

Finally, this study examined the degree to which these very early cognitive skills predicted a long-term adult outcome–years of schooling attained. As with the academic skills of the older children in the study, we predicted that any effect of the early cognitive skills on education attainment would be mediated by later academic skills. For working memory, the findings were consistent with our prediction. In contrast, attention at school entry predicted years of schooling attained, over and above academic achievement assessed in adolescence. Why might attention, as measured in this study, in early childhood predict persistence in school? We do not have a measure of attention beyond early adolescence, but it is possible that the ability to focus attention, inhibit impulses, and regulate activity level is somewhat stable, and continues to help people in achieving academic goals into adulthood. Also, attention is strongly associated with social skills as well as academic skills, and by helping children develop positive relationships with teachers and peers, attention skills may indirectly engender positive feelings about the school context, and thus a desire to remain in a school.

This finding suggesting possible long-term effects of attention is similar to a finding in the McClelland et al. (2013) study, in which attention span and persistence in tasks at age 4 years predicted the probability of completing college at age 25. Their measure of attention was different from the one used in the current study, but together the studies add to the growing evidence on the importance of attention-related skills in the early years of life.

Practical Implications

The finding that changes in attention and working memory predicted changes in academic achievement in the early but not the later grades suggest that beyond about fourth grade, efforts to improve academic achievement may be more productive if they are aimed directly at subject matter learning. In contrast, in the early grades children's academic achievement might be improved through interventions designed to develop attention and working memory. Little research exists to test this hypothesis, although it is developing. Current interventions for young children, such as Tools of the Mind (Barnett et al., 2008) and the Chicago School Readiness Project (CSRP; Raver et al., 2011), which were created to improve executive functions and self-regulation, provide some preliminary evidence on the potential value of efforts to improve underlying cognitive skills in the service of improving academic performance. There is experimental evidence using computer-based programs that working memory can be improved (Holmes, Gathercole, & Dunning, 2009; Kray & Ferdinand, 2013), but the effects of such interventions on children's math and literacy achievement have not been examined.

The strong associations between attention, memory and academic skills found in this study may have implications for how academic subjects are taught as well as for special intervention efforts. For example, math learning in young children might be enhanced by reducing the working memory demands in tasks or by teaching memory strategies in the context of teaching math. Performance in math and reading might also be improved by teaching children strategies for noticing when they become distracted. Evidence for the value of such efforts comes from research by Naglieri and Johnson (2000), in which children with ADHD performed better in math after they received direct instruction that included strategies to attend to relevant information and avoid distractions.

Limitations and Future Research

Although the results are statistically significant, the size of the effects are in many cases fractions of a standard deviation per grade. Certainly, there are other factors that could have an even larger influence on student academic trajectories. Nevertheless, cumulatively, over the course of 10 years, initial attention and memory scores predicted that students on average scored anywhere from 0.10 to 0.20 standard deviations higher academically than their peers who scored a standard deviation lower on these memory and attention measures initially. When one considers the size of this effect over time in relation to something like the Black-White achievement gap, which lingers around 0.70 standard deviations in kindergarten (Duncan & Magnuson, 2005), our findings suggested that targeted early childhood interventions to at-risk populations could make a significant dent.

Second, we did not assess each construct with multiple measures and correlations among measures were significant but modest. Some of our findings may be idiosyncratic to the measures used in this study (which were constrained by the NLSY data). It is possible that alternative measures of working memory or attention could demonstrate different associations with academic skills.

Third, while HLM's weighting and smoothing techniques adjust for missing outcome data at any given time point, it is still noteworthy that the NLSY suffered before the child supplement used in this study was implemented. Such minor nonrandom attrition could bias estimates if the important variables are not controlled for. However, data were never collected for less than 5% of children that are known to have been born to NLSY79 women. Further, an analysis by Aughinbaugh (2004) indicates that although children for whom supplemental information was not collected appear to be more disadvantaged than those for whom it was, there are very few such children; as a result their omission is likely to have a small effect on findings. Further, we control in our models for parent education, maternal age at birth of first child, and child race, which should adjust for important potential biases.

Finally, we note that our results are correlational in nature. The only compelling evidence of the causal direction between the kind of predictors examined in this study and academic achievement or attainment will come from intervention studies that examine the effects of promoting underlying cognitive skills, such as attention and working memory, on academic achievement, and the effects of improving academic achievement on these cognitive skills. For the future we highly recommend that such intervention studies include a broad array of outcome variables. Experimental or other studies designed to increase attention and memory need to include assessments of academic achievement, and interventions designed to improve academic achievement should include measures of these cognitive skills. Only this kind of intervention research will speak directly to issue of causality, and it has the added value of providing practical information on what kinds of interventions are effective.

References

Ackerman, P., Beier, M., & Boyle, M. (2005). Working memory and intelligence: The same or different constructs? *Psychological Bulletin, 131,* 30–60. http://dx.doi.org/10.1037/0033-2909.131.1.30

ACT. (2008). *The forgotten middle: Ensuring that all students are on target for college and career readiness before high school*. Iowa City, IA: Author.

Andersson, U. (2007). The contribution of working memory to children's mathematical word problem solving. *Applied Cognitive Psychology, 21,* 1201–1216. http://dx.doi.org/10.1002/acp.1317

Atkinson, R., & Shiffrin, R. (1968). Human memory: A proposed system and its control processes. In K. Spence & J. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 2, pp. 89–195). New York, NY: Academic Press. http://dx.doi.org/10.1016/S0079-7421(08)60422-3

Aughinbaugh, A. (2004). The impact of attrition on the children of the NLSY79. *The Journal of Human Resources, 39,* 536–563. http://dx.doi.org/10.2307/3559026

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. A. Bower (Ed.), *Recent advances in learning and motivation* (pp. 47–90). New York, NY: Academic Press.

Baker, P. (1993). *NLSY child handbook: A guide to the 1986–1990 National Longitudinal Survey of Youth Child Data, Issue 7610*. Columbus, OH: Center for Human Resource Research, Ohio State University.

Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychological Bulletin, 121,* 65–94. http://dx.doi.org/10.1037/0033-2909.121.1.65

Barnett, S., Jung, K., Yarosz, D., Thomas, J., Hornbeck, A., Stechuk, R., & Burns, S. (2008). Educational effects of the tools of the mind curriculum: A randomized trial. *Early Childhood Research Quarterly, 23,* 299–313. http://dx.doi.org/10.1016/j.ecresq.2008.03.001

Barriga, A. Q., Doran, J. W., Newell, S. B., Morrison, E. M., Barbetti, V., & Robbins, B. D. (2002). Relationships between problem behaviors and academic achievement in adolescents: The unique role of attention problems. *Journal of Emotional and Behavioral Disorders, 10,* 233–240. http://dx.doi.org/10.1177/10634266020100040501

Beck, S. J., Hanson, C. A., Puffenberger, S. S., Benninger, K. L., & Benninger, W. B. (2010). A controlled trial of working memory training for children and adolescents with ADHD. *Journal of Clinical Child and Adolescent Psychology, 39,* 825–836. http://dx.doi.org/10.1080/15374416.2010.517162

Berg, D. H. (2008). Working memory and arithmetic calculation in children: The contributory roles of processing speed, short-term memory, and reading. *Journal of Experimental Child Psychology, 99,* 288–308. http://dx.doi.org/10.1016/j.jecp.2007.12.002

Blair, C., Knipe, H., & Gamson, D. (2008). Is there a role for executive functions in the development of mathematics ability? *Mind, Brain, and Education, 2,* 80–89. http://dx.doi.org/10.1111/j.1751-228X.2008.00036.x

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

Booth, J., Boyle, J., & Kelly, S. (2010). Do tasks make a difference? Accounting for heterogeneity of performance of children with reading difficulties on tasks of executive function: Findings from a meta-analysis. *British Journal of Developmental Psychology, 28,* 133–176. http://dx.doi.org/10.1348/026151009X485432

Bull, R., Espy, K. A., & Wiebe, S. A. (2008). Short-term memory, working memory, and executive functioning in preschoolers: Longitudinal predictors of mathematical achievement at 7 years. *Developmental Neuropsychology, 33,* 205–228. http://dx.doi.org/10.1080/87565640801982312

Bull, R., & Scerif, G. (2001). Executive function as a predictor of children's mathematics ability: Inhibition, switching, and working memory. *Developmental Neuropsychology, 19,* 273–293. http://dx.doi.org/10.1207/S15326942DN1903_3

Butler, R., & Ruzany, N. (1993). Age and socialization effects on the development of social comparison motives and normative ability assessment in kibbutz and urban children. *Child Development, 64,* 532–543. http://dx.doi.org/10.2307/1131267

Davis, E., & Rowland, T. (1974). A replacement for the venerable Stanford-Binet? *Journal of Clinical Psychology, 30,* 517–521. http://dx.doi.org/10.1002/1097-4679(197410)30:4<517::AID-JCLP2270300418>3.0.CO;2-4

Diamantopoulou, S., Rydell, A. M., & Thorell, L. B. (2007). Impact of executive functioning and symptoms of attention deficit hyperactivity disorder on children's peer relations and school performance. *Developmental Neuropsychology, 32,* 521–542. http://dx.doi.org/10.1080/87565640701360981

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. http://dx.doi.org/10.1037/0012-1649.43.6.1428

Duncan, G. J., & Magnuson, K. A. (2005). Can family socioeconomic resources account for racial and ethnic test score gaps? *The Future of Children, 15,* 35–54. http://dx.doi.org/10.1353/foc.2005.0004

Dunn, L., & Dunn, L. (1981). *PPVT-R manual*. Circle Pines, MN: American Guidance Service.

Dunn, L., & Marwardt, F. (1970). *Peabody Individual Achievement Test manual*. Circle Pines, MN: American Guidance Service.

Engle, R. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science, 11,* 19–23. http://dx.doi.org/10.1111/1467-8721.00160

Engle, R., Tuholski, S., Laughlin, J., & Conway, A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General, 128,* 309–331. http://dx.doi.org/10.1037/0096-3445.128.3.309

Espy, K., McDiarmid, M., Cwik, M., Stalets, M., Hamby, A., & Senn, T. (2004). The contribution of executive functions to emergent mathematic skills in preschool children. *Developmental Neuropsychology, 26,* 465–486. http://dx.doi.org/10.1207/s15326942dn2601_6

Frick, P. J., Kamphaus, R. W., Lahey, B. B., Loeber, R., Christ, M. A. G., Hart, E. L., & Tannenbaum, L. E. (1991). Academic underachievement and the disruptive behavior disorders. *Journal of Consulting and Clinical Psychology, 59,* 289–294. http://dx.doi.org/10.1037/0022-006X.59.2.289

Gathercole, S. E., Alloway, T. P., Willis, C., & Adams, A. M. (2006). Working memory in children with reading disabilities. *Journal of Experimental Child Psychology, 93,* 265–281. http://dx.doi.org/10.1016/j.jecp.2005.08.003

Gathercole, S., Brown, L., & Pickering, S. (2003). Working memory assessments at school entry as longitudinal predictors of national curriculum attainment levels. *Educational and Child Psychology, 20,* 109–122.

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

Hinshaw, S. P. (1992). Externalizing behavior problems and academic underachievement in childhood and adolescence: Causal relationships and underlying mechanisms. *Psychological Bulletin, 111,* 127–155. http://dx.doi.org/10.1037/0033-2909.111.1.127

Hitch, G., Towse, J., & Hutton, U. (2001). What limits children's working memory span? Theoretical accounts and applications for scholastic development. *Journal of Experimental Psychology: General, 130,* 184–198. http://dx.doi.org/10.1037/0096-3445.130.2.184

Holmes, J., Gathercole, S. E., & Dunning, D. L. (2009). Adaptive training leads to sustained enhancement of poor working memory in children. *Developmental Science, 12,* F9–F15. http://dx.doi.org/10.1111/j.1467-7687.2009.00848.x

Klingberg, T. (2012). Is working memory capacity fixed? *Journal of Applied Research in Memory & Cognition, 1,* 194–196. http://dx.doi.org/10.1016/j.jarmac.2012.07.004

Kos, J. M., Richdale, A. L., & Hay, D. A. (2006). Children with attention deficit hyperactivity disorder and their teachers: A review of the literature. *International Journal of Disability, Development and Education, 53,* 147–160. http://dx.doi.org/10.1080/10349120600716125

Kray, J., & Ferdinand, N. (2013). How to improve cognitive control in development during childhood: Potentials and limits of cognitive interventions. *Child Development Perspectives, 7,* 121–125. http://dx.doi.org/10.1111/cdep.12027

Lan, X., Legare, C., Ponitz, C., Li, S., & Morrison, F. (2011). Investigating the links between the subcomponents of executive function and academic achievement: A cross-cultural analysis of Chinese and American preschoolers. *Journal of Experimental Child Psychology, 108,* 677–692. http://dx.doi.org/10.1016/j.jecp.2010.11.001

MacKinnon, D. P. (2008). *Introduction to statistical mediation analysis*. New York, NY: Taylor & Francis Group.

Markusic, M. (2012). *Understanding the Peabody Picture Vocabulary Test*. Troy, NY: Bright Hub Education. Retrieved from http://www.brighthub-education.com/special-ed-learning-disorders/13495-what-is-the-peabody-picture-vocabulary-test-all-about/

Martin, R., & Holbrook, J. (1985). Relationship of temperament characteristics to the academic achievement of first grade children. *Journal of Psychoeducational Assessment, 3,* 131–140. http://dx.doi.org/10.1177/073428298500300204

McClelland, M., Acock, A., Piccinin, A., Rhea, A., & Stallings, M. (2013). Relations between preschool attention span-persistence and age 25 educational outcomes. *Early Childhood Research Quarterly, 28,* 314–324. http://dx.doi.org/10.1016/j.ecresq.2012.07.008

Monette, S., Bigras, M., & Guay, M. (2011). The role of the executive functions in school achievement at the end of Grade 1. *Journal of Experimental Child Psychology, 109,* 158–173. http://dx.doi.org/10.1016/j.jecp.2011.01.008

Mott, F. (1995). *The NLSY Children1992: Description and evaluation*. Columbus, OH: Center for Human Resource Research, The Ohio State University. Retrieved from https://www.nlsinfo.org/sites/nlsinfo.org/files/attachments/121214/NLSYChildren1992Evaluation.pdf

Naglieri, J. A., & Johnson, D. (2000). Effectiveness of a cognitive strategy intervention to improve math calculation based on the PASS theory. *Journal of Learning Disabilities, 33,* 591–597. http://dx.doi.org/10.1177/002221940003300607

NICHD Early Child Care Research Network. (2003). Do children's attention processes mediate the link between family predictors and school readiness? *Developmental Psychology, 39,* 581–593. http://dx.doi.org/10.1037/0012-1649.39.3.581

Noël, M. (2009). Counting on working memory when learning to count and to add: A preschool study. *Developmental Psychology, 45,* 1630–1643. http://dx.doi.org/10.1037/a0016224

Oakhill, J. V., Cain, K., & Bryant, P. E. (2003). The dissociation of word reading and text comprehension: Evidence from component skills. *Language and Cognitive Processes, 18,* 443–468. http://dx.doi.org/10.1080/01690960344000008

Peterson, J. L., & Zill, N. (1986). Marital disruption, parent–child relationships, and behavior problems in children. *Journal of Marriage and the Family, 48,* 295–307. http://dx.doi.org/10.2307/352397

Rabe-Hesketh, S., & Skrondal, A. (2012). *Multilevel and longitudinal modeling using Stata: Vol. I. Continuous responses*. College Station, TX: Stata Press.

Raghubar, K., Barnes, M., & Hecht, S. (2010). Working memory and mathematics: A review of developmental, individual difference, and cognitive approaches. *Learning and Individual Differences, 20,* 110–122. http://dx.doi.org/10.1016/j.lindif.2009.10.005

Raudenbush, S. W., Bryk, A. S., Cheong, Y. F., Congdon, R. T., & du Toit, M. (2011). *HLM 7: Hierarchical linear and nonlinear modeling*. Lincolnwood, IL: Scientific Software International.

Raver, C. C., Jones, S. M., Li-Grining, C. P., Zhai, F., Bub, K., & Pressler, E. (2011). CSRP's impact on low-income preschoolers' pre-academic skills: Self-regulation and teacher-student relationships as two mediating mechanisms. *Child Development, 82,* 362–378. http://dx.doi.org/10.1111/j.1467-8624.2010.01561.x

Roid, G. H. (2003). *Stanford-Binet Intelligence Scales* (5th ed.). Itasca, IL: Riverside.

Rothbart, M. K., & Jones, L. B. (1998). Temperament, self-regulation, and education. *School Psychology Review, 27,* 479–491.

Ruble, D., Grosovsky, E., Frey, K., & Cohen, R. (1992). Developmental changes in competence assessment. In A. Boggiano & T. Pittman (Eds.), *Achievement and motivation: A social-developmental perspective* (pp. 138–164). Cambridge, England: Cambridge University Press.

Rudasill, K., Gallagher, K., & White, J. (2010). Temperamental attention and activity, classroom emotional support, and academic achievement in third grade. *Journal of School Psychology, 48,* 113–134.

Savage, R. S., Frederickson, N., Goodwin, R., Patni, U., Smith, N., & Tuersley, L. (2005). The relationship among rapid digit naming, phonological processing, motor automaticity and speech perception in poor average and good readers and spellers. *Journal of Learning Disabilities, 38,* 12–28. http://dx.doi.org/10.1177/00222194050380010201

Savage, R., Lavers, N., & Pillay, V. (2007). Working memory and reading difficulties: What we know and what we don't know about the relationship. *Educational Psychology Review, 19,* 185–221. http://dx.doi.org/10.1007/s10648-006-9024-1

Shipstead, Z., Hicks, K., & Engle, R. (2012). Cogmed working memory training: Does the evidence support the claims? *Journal of Applied Research in Memory & Cognition, 1,* 185–193. http://dx.doi.org/10.1016/j.jarmac.2012.06.003

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York, NY: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780195152968.001.0001

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

Stepankova, H., Lukavsky, J., Buschkuehl, M., Kopecek, M., Ripova, D., & Jaeggi, S. (2014). The malleability of working memory and visuospatial skills: A randomized controlled study in older adults. *Developmental Psychology, 50,* 1049–1059.

Stipek, D. (2002). *Motivation to learn: Integrating theory and practice* (4th ed.). Needham Heights, MA: Allyn & Bacon.

Swanson, H. L., & Beebe-Frankenberger, M. (2004). The relationship between working memory and mathematical problem solving in children at risk and not at risk for serious math difficulties. *Journal of Educational Psychology, 96,* 471–491. http://dx.doi.org/10.1037/0022-0663.96.3.471

Swanson, H. L., & Jerman, O. (2006). Math disabilities: A selective meta-analysis of the literature. *Review of Educational Research, 76,* 249–274. http://dx.doi.org/10.3102/00346543076002249

Swanson, H. L., & Jerman, O. (2007). The influence of working memory on reading growth in subgroups of children with reading disabilities. *Journal of Experimental Child Psychology, 96,* 249–283. http://dx.doi.org/10.1016/j.jecp.2006.12.004

Terman, L. M. (1916). *The measurement of intelligence*. Boston, MA: Houghton Mifflin. http://dx.doi.org/10.1037/10014-000

Thompson, R. A., & Nelson, C. A. (2001). Developmental science and the media: Early brain development. *American Psychologist, 56,* 5–15. http://dx.doi.org/10.1037/0003-066X.56.1.5

Wechsler, D. (1974). *Manual for the Wechsler Intelligence Scales for Children*. New York, NY: The Psychological Corporation.

Welsh, J. A., Nix, R., Blair, C., Bierman, K., & Nelson, K. (2010). The development of cognitive skills and gains in academic school readiness for children from low-income families. *Journal of Educational Psychology, 102,* 43–53. http://dx.doi.org/10.1037/a0016738

Wigfield, A., Eccles, J., Yoon, K., Harold, R., Arbreton, A., Dreedman-Doan, C., & Blumenfeld, P. (1997). Changes in children's competence beliefs and subjective task values across the elementary school years: A three-year study. *Journal of Educational Psychology, 89,* 451–469. http://dx.doi.org/10.1037/0022-0663.89.3.451

Williams, P., Weiss, L., & Rolfhus, E. (2003). *WISC technical report #2: Psychometric properties*. London, United Kingdom: Psychological Corporation. Retrieved from http://faculty.utep.edu/Portals/1669/Shu/5335/stanford_binet.pdf

Zill, N. (1990). *Behavior Problems Index Based on Parent Report* (Publication no. 1991–03). Bethesda, MD: Child Trends.

Received February 1, 2014
Revision received August 11, 2014
Accepted August 15, 2014 ■