

Schooling Effects on Literacy Skills During the Transition to School

Matthew H. Kim 

University of Washington

Frederick J. Morrison

University of Michigan

The current study used a regression discontinuity (RD) design to characterize more precisely the link between schooling and literacy by examining whether and how different grade-level, practice-as-usual schooling experiences uniquely predict specific literacy subskills during the transition to school. Data from 334 children revealed moderate positive effects of pre-kindergarten, kindergarten, and first grade schooling on decoding, while kindergarten and first grade schooling predicted comprehension skills. There was no significant effect of schooling at any grade level on expressive vocabulary or sound awareness. Results were robust to different RD estimation methods and highlight the heterogeneity of schooling effects on early literacy skill development. Implications for understanding early literacy development in the context of regular, public schooling are discussed.

Keywords: *literacy, decoding, comprehension, early schooling, regression discontinuity*

Introduction

LEARNING to read is among the most important skills that children acquire during the early childhood years. Reading development is a complex and multifaceted process and involves the acquisition and mastery of many subskills, including decoding, sound awareness, vocabulary, and comprehension. While these skills grow as a function of biological development, children also acquire these skills through multiple, extensive experiences at home and especially in school. Research has demonstrated that special intervention programs positively contribute to literacy skills, but less is known about whether the effects of practice-as-usual literacy instruction vary as a function of grade-level schooling experiences and different literacy skills important for fluent reading. In the present study, we aim to characterize more precisely the link between schooling and literacy by examining whether and how different grade-level schooling experiences uniquely predict specific literacy subskills. Toward that end, we used a quasi-experimental method called regression discontinuity to examine the nature and magnitude of schooling effects on a range of skills representative of early literacy during the transition to school.

Development of Early Reading

Fluent reading involves mastery of a variety of literacy skills during early childhood and elementary school. Whitehurst and Lonigan (1998) describe *emergent literacy* as

a developmental phenomenon in which meaningful reading skill acquisition occurs during the preschool years, contrasting with the traditional view that reading skills are largely learned only when children start formal schooling. Reading development is shaped by a variety of individual and contextual factors, including the home and parenting environment (e.g., Roberts, Jurgens, & Burchinal, 2005). Individual differences in literacy behaviors emerge very early in life and remain stable throughout early childhood and even beyond (e.g., Hart & Risley, 1995), but growth in literacy is also dependent on schooling and can be modified through regular instruction as well as special intervention programs (e.g., Borman et al., 2007; Farver, Lonigan, & Eppe, 2009; Schwartz, 2005; Wilson, Dickinson, & Rowe, 2013).

The current study examined four domains of early literacy skill development: decoding, phonological skills, vocabulary, and comprehension. Decoding involves the ability to map alphabetic symbols and words with their corresponding sounds (e.g., Lonigan, Burgess, & Anthony, 2000). Phonological skills reflect the ability to identify, analyze, and manipulate sounds in language and include knowledge of rhyme and syllables (e.g., Whitehurst & Lonigan, 1998). Vocabulary, or word learning, is often conceptualized as having two components—receptive vocabulary (recognizing words) and expressive vocabulary (recalling words)—and develops rapidly during the preschool years (e.g., Sénéchal, 1997). Finally, comprehension, itself a multifaceted construct that consists of multiple cognitive processes



and develops in interaction with instruction, is critical for fluent reading (e.g., Connor, 2016; Keenan, Betjemann, & Olson, 2008).

Of particular relevance is the body of evidence demonstrating that different literacy skills exhibit different developmental trajectories (e.g., Dickinson, McCabe, Anastasopoulos, Peisner-Feinberg, & Poe, 2003; Kendeou, van den Broek, White, & Lynch, 2009; Lonigan et al., 2000; National Early Literacy Panel (NELP), 2008), consistent with the established perspective within the field that early literacy skills continuously develop during early childhood. For example, phonological awareness exhibits a developmental continuum that increases in linguistic complexity, from word and syllable awareness to onset-rime and phonemic awareness (Phillips, Clancy-Menchetti, & Lonigan, 2008). What is less well understood is the proper timing and delivery of instructional practices for targeting different literacy skill sets. Because literacy development often occurs in stages, one potential strategy might be to focus a greater proportion of instructional time on basic, code-focused skills in the early grades before switching emphasis to more advanced, meaning-focused skills in later grades. However, because basic skills continue to develop concurrently with the acquisition of more complex reading skills, this strategy oversimplifies the complexity of literacy development and overlooks the significant individual differences in how children acquire literacy skills.

In summary, while the idea of providing age-appropriate instruction in literacy skills is a popular one, the field still lacks sufficient evidence regarding how children of different ages and in different grades experience literacy instruction and subsequent literacy outcomes. Indeed, one of the recommendations of the National Early Literacy Panel (NELP) is to conduct direct tests of age differentiation in early literacy instruction across preschool and kindergarten (NELP, 2008). The present study seeks to answer this call by examining whether different grade-level schooling experiences—over and above the effects of age—lead to gains in different literacy skills during the school transition period.

Schooling and Early Literacy

Reading is a dominant focus of academic instruction during the early grades. In prekindergarten and kindergarten classrooms, more time is spent on literacy-related activities than on other academic domains (e.g., Kim, Bell, & Morrison, 2011; Phillips, Gormley, & Lowenstein, 2009). Recent research has revealed a distinct shift toward more formal instruction in earlier grades, with one recent piece suggesting that kindergarten is the new first grade (Bassok, Latham, & Rorem, 2016). This shift has sparked a growing interest in the nature of schooling during the transition to school as well as the unique impact of schooling on early academic skills, particularly literacy. A better understanding

of which literacy skills are most impacted by schooling—and at which particular grade levels—would provide important insights regarding the nature of reading development during the school transition period and point to potential avenues for intervention and improvement.

Evidence regarding the impact of schooling on early literacy are primarily drawn from two types of early schooling experiences: (1) literacy intervention programs and (2) practice-as-usual literacy instruction. This distinction is important for our understanding of the trajectories and mechanisms of literacy skill acquisition, as the method, nature, and timing of literacy instruction in schools can provide insights into effective instructional practices in the early grades.

Literacy Intervention Programs. Consistent with work suggesting that relatively intensive and multifaceted interventions have the greatest potential to produce maximal literacy outcomes (Whitehurst & Lonigan, 1998), a strong body of empirical evidence demonstrates that such intervention programs indeed improve early literacy outcomes (e.g., Borman et al., 2007; Schwartz, 2005; Wilson et al., 2013). These programs are distinct from core reading curricula in that teachers often receive coaching and specialized professional development. In addition, these interventions are traditionally implemented on a smaller scale and often target subgroups of students with particular learning needs rather than an entire classroom. However, examining schooling effects based on literacy intervention programs is likely not representative of the majority of classrooms, which do not adopt a special intervention program to deliver reading instruction to all students, or lack the resources or supports to ensure that the program is implemented with a high degree of fidelity. Furthermore, interventions that produced the largest positive effects on children's early literacy skills were often conducted in one-on-one or small-group instructional activities (NELP, 2008), highlighting the need to better understand the effects of conventional literacy instruction in general education classrooms in larger group settings.

Practice-As-Usual Instruction. In contrast to specialized intervention programs that offer intensive supports to teachers, most classrooms adopt a standard core reading curriculum intended to meet the needs of all students. In implementing these standard curricula, teachers often have wide latitude in how they actually deliver literacy instruction in their classrooms and typically do not receive the intensive supports that teachers implementing more specialized programs often receive. Examining the effects of practice-as-usual instruction is critical to understanding whether and how “regular” schooling—not just specialized programs—influences early literacy outcomes.

Exploring practice-as-usual effects has proven challenging for researchers because schooling is conflated with age-related changes among students, making it difficult to

pinpoint whether growth in academic skills is due to unique schooling experiences or maturational and other environmental causes. Because education is compulsory in the United States, it is impossible to randomly assign same-aged children to a schooling group and a non-schooling group. Whereas studies of intervention programs often use experimental designs that randomly assign classrooms to treatment and control conditions, studies of practice-as-usual effects often rely on quasi-experimental techniques that can generate relatively unbiased estimates of the unique impact of schooling, over and above the effect of age, allowing researchers to make causal statements regarding the effect of schooling on a range of child outcomes. Put more simply, examining practice-as-usual effects is equivalent to investigating whether children who receive typical (non-intervention) literacy classroom instruction experience better outcomes compared to same-aged children who do not receive that instruction.

School cutoff and regression discontinuity are two approaches that have been commonly used to isolate the effects of schooling on early literacy. In the school cutoff design (e.g., Morrison, Smith, & Dow-Ehrensberger, 1995), researchers capitalize on the fact that many school districts have a cutoff date for school enrollment, such that children born on or before a given date can enter school (usually kindergarten) for that academic year, but children born after that date must wait until the following year. Based on this cutoff date, children are separated into two groups—one that attends kindergarten and one that does not. Regression discontinuity (RD) offers a more flexible approach to the school cutoff technique. In RD, children are also located at different points on the date-of-birth continuum, with the cutoff date for school enrollment separating the two groups. A best-fit regression line is then fit through the data on either side of the cutoff, with a “jump” or discontinuity at the cutoff indicating that the difference in outcomes between the two groups is likely to be due to schooling. Therefore, in both designs, the “treatment” is receiving standard literacy instruction in kindergarten (or first grade), while the “control” or “counterfactual” is receiving standard literacy instruction in prekindergarten (or kindergarten).

Research examining practice-as-usual literacy instruction has revealed schooling effects on many, but not all, early literacy skills. In one of the pioneering studies using a quasi-experimental design, Cahan and Cohen (1989) found that schooling had larger effects on verbal skills compared to non-verbal skills, and the effect of one year of schooling was larger than the effect of one year of age. Recently, studies using the school cutoff design have shown positive prekindergarten effects (Burrage et al., 2008) and kindergarten effects (Burrage et al., 2008; Christian, Morrison, Frazier, & Massetti, 2000) on decoding skills, as well as first grade effects on phonemic segmentation skills (Christian et al., 2000; Morrison et al., 1995). Studies of state pre-K programs

using an RD design have also shown positive effects on phonemic segmentation (Gormley, Gayer, Phillips, & Dawson, 2005; Weiland & Yoshikawa, 2013). Evidence for schooling effects on vocabulary development is mixed, with one study showing a pre-K effect on receptive vocabulary (Weiland & Yoshikawa, 2013), but other investigations demonstrating that receptive vocabulary (Christian et al., 2000) and expressive vocabulary (Skibbe, Connor, Morrison, & Jewkes, 2011) are more sensitive to age-related effects. These mixed findings might be attributable to a lack of consistent and explicit vocabulary instruction in pre-K and kindergarten (Juel, Biancarosa, Coker, & Deffes, 2003; Neuman & Dwyer, 2009) or because family-level characteristics exert stronger influences than schooling effects on vocabulary outcomes at this age (Connor, Son, Hindman, & Morrison, 2005), or some combination of these and other factors. Another study demonstrated both age-related and schooling-related change in story recall and story production, important prerequisite skills for comprehension (Varnhagen, Morrison, & Everall, 1994). Schooling effects have also emerged in other literacy skills including spelling (Gormley et al., 2005) and print knowledge (Wong, Cook, Barnett, & Jung, 2008).

Contributions of the Present Study

While these studies reveal that schooling is important for developing literacy skills, much of what we know regarding the effects of practice-as-usual instruction on early literacy development come from individual studies—often using cross-sectional data—that have either examined a schooling effect at a single grade level or have examined just one or two literacy skills. While longitudinal designs can provide useful insights regarding patterns of growth, they are not appropriate for drawing causal inferences regarding the unique effect of schooling on academic outcomes, over and above age-related influences. Recent methodological advances in RD have made it possible for researchers to implement this design in broader contexts using increasingly sophisticated estimation methods. Accordingly, the present study examines longitudinal data using an RD design to conduct a rigorous analysis of multiple grade-level, practice-as-usual schooling effects on a range of literacy skills that more fully reflect the complexity of early literacy development, allowing us to characterize more precisely the link between schooling and literacy.

Data examined in the present study have been previously analyzed using other methods, including the school cutoff design and nonlinear growth curve modeling (e.g., Burrage et al., 2008; Skibbe et al., 2011; Skibbe, Grimm, Bowles, & Morrison, 2012; Skibbe, Montroy, Bowles, & Morrison, 2018). Recognizing the increasing awareness and importance of replicability in the education sciences, the present study seeks not only to replicate previous findings of schooling effects on literacy skills, but to extend this work by

applying recent advances in the age-cutoff RD design in order to yield new insights into the developmental trajectories of the components of emerging literacy and, in particular, whether and how schooling experiences influence early literacy outcomes.

Study Aim and Hypotheses

The present study seeks to examine the unique effect of schooling at *three* grade levels—prekindergarten, kindergarten, and first grade—on each of *four* distinct skills representative of early literacy—decoding, sound awareness, expressive vocabulary, and comprehension—using an age-cutoff RD design. Our hypotheses are guided both by our knowledge of the emergence and developmental trajectories of various literacy skills in early childhood, as well as recent findings from quasi-experimental studies of schooling and literacy. Given previous research, we expected basic, code-focused skills such as decoding and sound awareness to have an earlier developmental trajectory, while more advanced, meaning-focused skills such as vocabulary and comprehension would emerge later. Accordingly, while we predicted that schooling would be positively related to children’s decoding, sound awareness, and comprehension at each grade level, we also expected a stronger effect of pre-K and kindergarten on decoding skills and sound awareness compared to first grade. In contrast, we expected a stronger effect of first grade on comprehension compared to pre-K and kindergarten. However, we predicted that schooling would not impact children’s expressive vocabulary at any grade level, given the predominance of evidence indicating that expressive vocabulary is predicted by age-related change rather than schooling.

Method

Participants

Three hundred thirty-four children ($M = 5.64$ years, $SD = 0.93$, Range = 3.75–7.67, 174 girls) were assessed as part of a longitudinal study of literacy development. Children were recruited from 16 schools in a single school district in the Midwestern United States. At the time of assessment, the district adopted a standard, general education reading curriculum called Open Court, and literacy instruction in these classrooms was considered to be practice-as-usual. Typical of core reading curricula, Open Court introduces and reinforces a variety of literacy skills at each grade, but the amount and type of instruction devoted to each skill changes somewhat over time as a function of children’s developing knowledge and readiness to acquire new skills. For example, letter recognition and phonemic awareness are emphasized in earlier grades, while greater attention is devoted to vocabulary and comprehension as children progress through elementary school. Nevertheless, basic skills continue to be

reinforced in later grades, while advanced skills such as comprehension are introduced and taught as early as kindergarten, with implementation of instruction depending partly on the teacher and classroom composition. Previously published empirical work using data from the same study indicates that preschool classrooms in the district followed the same academic calendar as elementary schools and did not offer summer programming, and that participating children did not meaningfully differ from children whose parents did not return a consent form (Connor, Morrison, & Slominski, 2006; Skibbe et al., 2012). The statewide cutoff date for school entry was December 1, so children who were five years of age on or before December 1 were eligible to enter school that academic year, while children who were not yet five years old on December 1 had to delay enrollment until the following year. Seventy-eight percent of children were White/Caucasian, 79.5% of mothers had at least a college degree, and the median household income was \$119,000, reflecting a relatively racially homogeneous, middle- to upper-middle socioeconomic status (SES) sample.

Procedure

Parental written consent was obtained and child oral assent was received prior to data collection. Because of the structure of recruitment and assessment of children, which occurred on a rolling basis during a five-year period, two cohorts of data were available for each grade during years two, three, and four of the study. For example, pre-K children were recruited and assessed in Year 2 (Cohort 1), and a separate cohort of pre-K children were recruited and assessed in Year 3 (Cohort 2). Data were pooled across cohorts, as there were no significant differences between the two cohorts on key baseline characteristics. (As described in Skibbe et al. (2012), efforts to minimize attrition included the recruitment of additional four-year-olds (pre-K children) during Year 2 and five-year-olds (kindergarten children) during Year 3.) Because children were assessed in the fall and spring of each study year, it was possible to estimate three different schooling effects—pre-K, kindergarten, and first grade. Importantly, the timing and duration of fall and spring data collection periods were similar across study years, reducing the possibility that any observed schooling effects might be due to differences in the timing of assessment (see Table 2 in Skibbe et al., 2012, for information on average date of testing for each time point). Detailed information on how the data set was constructed for use in the RD analysis is presented in the *Data Structure* subsection.

Measures

Measures of early literacy were drawn from the Woodcock Johnson (WJ) Tests of Achievement III (Woodcock, McGrew, & Mather, 2001), a nationally normed and validated set of assessments that have been widely and successfully used in

TABLE 1
Descriptives for Woodcock Johnson Outcome Variables

Variables	Mean	SD	Range
<i>Prekindergarten</i>			
Letter-word identification	344.61 (116.57)	25.94 (16.31)	270–468
Sound awareness	447.75 (123.87)	17.93 (16.00)	420–492
Picture vocabulary	472.33 (112.35)	11.06 (10.86)	439–502
Passage comprehension	401.13 (118.02)	15.35 (10.93)	350–477
<i>Kindergarten</i>			
Letter-word identification	395.13 (115.42)	43.46 (13.68)	270–528
Sound awareness	475.81 (121.53)	19.11 (14.69)	420–529
Picture vocabulary	482.56 (112.35)	11.90 (11.06)	435–528
Passage comprehension	425.81 (108.54)	31.06 (14.96)	358–512
<i>First grade</i>			
Letter-word identification	458.93 (117.24)	27.55 (11.66)	389–528
Sound awareness	493.16 (116.42)	12.63 (13.51)	461–529
Picture vocabulary	489.98 (111.14)	10.29 (10.38)	456–528
Passage comprehension	470.37 (109.83)	16.09 (11.23)	391–515

Descriptives represent W scores (with standard score in parentheses; $M = 100$, $SD = 15$). W score is on an equal-interval scale.

this age range. Children were assessed on four subtests, each reflecting a different aspect of early reading. The Letter-Word Identification subtest (decoding) required children to identify and pronounce letters and words. The Sound Awareness subtest (sound awareness) assessed skills related to the rhyming, deletion, substitution, and reversal of syllables and phonemes. The Picture Vocabulary subtest (expressive vocabulary) required children to name objects represented by pictures. The Passage Comprehension subtest (symbolic understanding and comprehension) required children to match words with pictures, identify words corresponding to phrases, and identify words that complete a sentence or passage. The WJ is specifically designed to generate scores that are comparable across ages, and empirical work has demonstrated that the basal and ceiling levels that determine the start and end rules are correctly defined (Watts, Spanier, & Duncan, 2014).

W scores were obtained using the WJ-III Compuscore software program. The W score metric is on an equal-interval scale and allows us to directly compare the achievement of one student against another, regardless of age (Jaffe, 2009). Sample means for the WJ literacy subtests by grade level are presented in Table 1. The corresponding standard score means are larger than 100, indicating that our sample is characterized as relatively high achieving, but there still exists considerable variability in these means. Bivariate correlations between the WJ literacy subtests are presented in Table 2. While there are strong positive correlations between the subtests, the coefficients are not close to 1, indicating that these literacy skills are distinct from each other.

RD Specification

The utility of the RD technique to generate relatively unbiased estimates of a treatment—even in the absence of randomization—comes with a set of potential threats to internal validity that must be addressed before valid causal inferences are made. RD can be implemented in a variety of ways, informed by theory as well as the nature and structure of the data. Given the relatively small sample size in the present study, we purposefully adopted a conservative analytical approach while running a series of robustness checks to supplement the primary analysis. A number of guides and empirical papers provide an excellent, detailed explanation of the technique as well as its implementation (e.g., Bloom, 2009; Imbens & Lemieux, 2008; Jacob, Zhu, Somers, & Bloom, 2012; Lipsey, Weiland, Yoshikawa, Wilson, & Hofer, 2015). This section provides details on the major features of the RD implementation in the current study, which are based on previous empirical examples as well as best practices from the psychology and econometrics literature.

Running Variable. Assignment to treatment (i.e., schooling) is determined by an individual’s position on a continuous scale relative to a cutoff point (i.e., cutoff date for school enrollment based on a child’s date of birth); this scale is called the running variable (also known as a forcing variable or a quantitative assignment variable). Consistent with previous investigations, we allowed each child to have an integer value on our *days from cutoff* running variable, where negative values indicated that the child missed the cutoff for school entry, and zero and positive values indicated that the child made the cutoff.

TABLE 2

Bivariate Correlations for Woodcock Johnson Outcome Variables

Variables	LW	SA	PV	PC
Letter-word identification (LW)	1.00			
Sound awareness (SA)	.79***	1.00		
Picture vocabulary (PV)	.89***	.72***	1.00	
Passage comprehension (PC)	.58***	.64***	.55***	1.00

Coefficients represent Pearson correlations using listwise deletion.

*** $p < .001$.

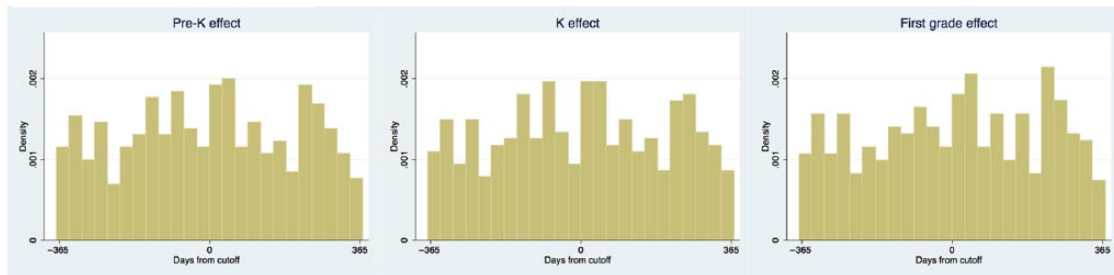


FIGURE 1. *Histograms showing the density of values of the days from cutoff running variable.*

Each bin represents approximately 30 days (one month). There is no evidence of treatment manipulation in the children’s enrollment status at the cutoff (i.e., where days from cutoff equals zero).

We would expect that children’s birth dates are equally distributed across the calendar year. However, if parents make conscious decisions to ensure that their child is born before (or after) the school enrollment cutoff date, we would observe clustering of birth dates on one side of the cutoff, which would potentially threaten the internal validity of the design. In order to rule out this potential confound, we generated histograms centered on the DOB (date of birth) cutoff for each schooling effect in order to examine the density of data around the cutoff. As shown in Figure 1, there is no evidence indicating an abnormal cluster of birth dates on one side of the cutoff or at any other time point for each of the three schooling effects.

Functional Form. Functional form is a term that describes the true underlying relation between the running variable and the outcome. We assumed that the relation between child age and literacy skills in our relatively protracted time period was linear. Consistent with best practice, we also generated quadratic and local linear estimates as a robustness check, which allow for more flexibility in modeling the relation between the running variable and the outcome. While cubic and higher-order polynomials are sometimes used in RD, these polynomials might introduce unnecessary noise in the estimates that lead to potentially unreliable estimates (Gelman & Imbens, 2014).

Covariates. One potential threat to internal validity is the possibility that any apparent schooling effect might be due

not to the treatment, but to some other variable. For example, a discontinuity at the cutoff might not be due to schooling, but rather because children who made the cutoff for school entry also had parents who were more highly educated compared to children who missed the cutoff. Recent advances in the RD technique have focused on the appropriateness of including covariates in generating treatment estimates. Research indicates that including covariates can improve precision in the estimates but that they are not required for obtaining unbiased or consistent estimates (Jacob et al., 2012). Similarly, under minimal smoothness assumptions (described in the following paragraph), covariates can improve precision while point estimates remain stable (Calonico, Cattaneo, Farrell, & Titiunik, 2016).

The selection of covariates depends not only on the theoretical link between each covariate and the outcome variable, but also practical considerations associated with preserving degrees of freedom when sample size is limited. Our approach was to include a small number of baseline characteristics related to achievement—child gender, maternal education, and household income—in order to improve the precision of our RD estimates. Given the limited geographical region of the schools from which students were recruited (i.e., a single school district), we chose not to control for community-level characteristics. We then examined the degree to which key baseline characteristics were “smooth through the threshold” to ensure that these other variables were not driving any observed effect. As shown in Table S1 (Online Supplemental Material), the two groups

TABLE 3
Analytic Sample Sizes by Schooling Effect and Bandwidth

Variables	12 months		10 months		8 months		6 months		4 months	
Pre-K effect	420		350		281		226		154	
<i>(missed cutoff / made cutoff)</i>	205	215	165	185	136	145	113	113	71	83
Kindergarten effect	411		342		274		220		149	
<i>(missed cutoff / made cutoff)</i>	198	213	160	182	132	142	108	112	67	82
First grade effect	392		326		258		203		141	
<i>(missed cutoff / made cutoff)</i>	184	208	147	179	118	140	95	108	64	77

Months represent bandwidths. Preferred bandwidth is six months (180 days). Only children with complete data on all variables are counted. For bandwidths greater than six months, some children are located on both sides of the cutoff and are thus double counted.

were not significantly different at the cutoff on child age, child gender, maternal education, and household income.

Finally, because children were drawn from 16 schools, it is important to assess the degree to which school-level differences might be driving any observed schooling effect. For example, if a school or set of schools is overrepresented on one side of the cutoff, any effect that we observe might be due to particular features of specific schools, rather than our hypothesis that schooling in general is a unique predictor of children’s literacy growth. Given the lack of detailed information available on each of the schools, controlling for possible school-level differences is especially important. We found some evidence that certain schools were overrepresented on one side of the cutoff. Because of this, we included school fixed effects as additional covariates. (Because of the sheer number of discrete classroom units to which children were assigned, as well as our relatively small sample size, we decided to control for variance at the school level rather than at the classroom level in order to preserve as many degrees of freedom as possible.) As a robustness check, we generated estimates based on a restricted sample of children who were enrolled only in schools that appear on both sides of the cutoff, as shown in Table S2 (Online Supplemental Material); estimates on the restricted sample are virtually identical to those on the full sample.

Bandwidth. Bandwidth is the window around the cutoff that is used to generate estimates of the treatment. In general, the narrowest possible bandwidth is preferred, but the choice of bandwidth ultimately reflects a tradeoff between bias and variance. Comparing outcomes between children born on November 30 with children born on December 2 would lead to the most accurate estimates—as the children are virtually the same age—but the extremely small number of children born in such a narrow DOB window would lead to noisy estimates. On the other hand, a wider window would reduce noise because more observations are being used to generate the regression estimates, but children far from the cutoff are being used to inform the jump at the cutoff, which would lead to less accurate estimates.

Previous empirical work on schooling effects on literacy outcomes have used a six-month bandwidth either as a primary specification or as a robustness check (Gormley et al., 2005; Weiland & Yoshikawa, 2013; Wong et al., 2008). However, given the small sample size in the present study, adopting a 12-month (365 days) bandwidth would allow us to retain the maximum number of observations for the analysis. Recognizing this tradeoff, we adopted a six-month (180 days) bandwidth in order to maximize the comparability of our results across similar studies, and then assessed the degree to which our estimates were sensitive to bandwidths of four, eight, ten, and 12 months. For bandwidths greater than six months, it is necessary to cluster standard errors by child in order to account for the fact that some children will appear on both sides of the cutoff; we accomplished this by generating robust standard errors. Figure S1 (Online Supplemental Material) shows that our estimates are robust to different bandwidths. When calculating local linear estimates, a data-driven cross-validation method was used to determine the optimal bandwidth, according to the method described in Calonico, Cattaneo, and Titiunik (2014). Table 3 presents analytic sample sizes by schooling effect and bandwidth.

Data Structure. Because of the longitudinal nature of the study, data from the same children were used to calculate each grade-level schooling effect. This allowed us to control for variability within individuals while simultaneously conducting a between-subjects RD analysis. The pre-K effect was assessed by comparing the outcomes between pre-K children and kindergarten children assessed in the fall of the school year. By comparing the outcomes between kindergarten children who just made the cutoff and pre-K children who just missed the cutoff, the fall assessment for the kindergarten children reflects the schooling they received the previous year—that is, pre-K. The kindergarten effect was assessed by comparing the outcomes between the same children assessed in the *spring* of the school year, after kindergarten children had experienced a full year of kindergarten. Finally, the first grade effect was assessed by comparing the

TABLE 4

Cohort Data Collection and Construction of Regression Discontinuity Models

Comparison	Pre-K effect	Kindergarten (K) effect	1st grade effect
	RD models 1a–1d	RD models 2a–2d	RD models 3a–3d
Missed cutoff	Pre-K (Y2, fall) + Pre-K (Y3, fall)	Pre-K (Y2, spring) + Pre-K (Y3, spring)	K (Y3, spring) + K (Y4, spring)
Made cutoff	K (Y2, fall) + K (Y3, fall)	K (Y2, spring) + K (Y3, spring)	1st (Y3, spring) + 1st (Y4, spring)

Data from study years two through four were used in the present analysis, as there were two cohorts of children available in each of those years. Two cohorts of children were combined to calculate each grade-level effect. Within each grade-level effect, there were a small number of children who appeared on both sides of the cutoff. For example, in the estimation of the pre-K effect, some pre-K children assessed in the fall of Y2 were kindergarten children assessed in the fall of Y3. However, by constraining our bandwidth to six months in our primary RD specification, no children appeared on both sides of the cutoff.

outcomes between kindergarten children and first grade children assessed in the spring of the school year. Table 4 presents information on how the RD models were constructed to assess each of the three grade-level schooling effects.

While it was possible to measure the kindergarten effect by comparing the outcomes between kindergarten and first grade children assessed in the fall of the school year, we chose not to do so for three reasons. First, assessing children in the spring reduced the impact of attrition between the end of one school year and the beginning of the next school year, particularly important in the context of this small-scale RD design. Second, because we assessed the first grade effect by examining spring scores, assessing the kindergarten effect by also examining spring scores permits a more direct comparison. Third, it is possible that some families in our sample sought out and enrolled their children in summer activities. Participation in these enrichment programs could have mitigated fadeout effects for some children but not for others. Because we did not have access to data on children’s participation in summer activities, calculating schooling effects using spring scores for these children eliminated the potential confound of the subsequent summer months.

Treatment Misallocation. While children may be assigned to a treatment based on their location on the running variable, these same children may not actually receive the treatment for a number of reasons. For example, children born on November 30—while eligible to enroll in school during that academic year—may not actually enroll due to parent concerns that the child would be the youngest student in the classroom, the child might not be behaviorally or emotionally ready for school, or some other reason. This common phenomenon is known as *redshirting* and is likely to occur for children who just make the cutoff for school entry (e.g., Bassok & Reardon, 2013; Lincove & Painter, 2006). While less common, children who are not eligible for school entry based on their DOB might end up enrolling in school.

These “crossover” children reflect a potential issue known as *treatment misallocation*, where the child’s expected treatment status is not identical to the child’s actual

treatment status (in our sample, noncompliance rates ranged from 4.8 to 5.1% across the entire sample, comparable with previous studies). These cases reflect what is commonly known as a *fuzzy RD*, and different estimation methods are available that account for misallocation. In a treatment on the treated (TOT) analysis, the estimates of the jump at the cutoff are rescaled to account for this misallocation, such that the resulting estimates reflect the treatment effect on treated individuals. In an intent-to-treat analysis (ITT), no rescaling occurs; rather, these estimates reflect the treatment effect while ignoring this misallocation. Accordingly, ITT estimates tend to be more conservative and reflect the actual implementation of the school enrollment cutoff in the district. Previous empirical studies have adopted a variety of different approaches, with some focusing on TOT (e.g., McEwan & Shapiro, 2008), others that exclude crossovers entirely (e.g., Gormley et al., 2005), and still others that present results from multiple estimation methods (e.g., Weiland & Yoshikawa, 2013; Wong et al., 2008).

Estimation Method. Given our relatively small sample size, our preferred approach is to highlight the more conservative reduced-form ITT estimates. We estimate the following regression:

$$outcome_{ij} = \alpha_0 + \alpha_1 dfc_{ij} + \alpha_2 Tx_{ij}^{exp} + \alpha_3 dfc_{ij} \times Tx_{ij}^{exp} + \alpha_4 covariates_{ij} + \alpha_5 sfe_j + \varepsilon_{ij}$$

where $outcome_{ij}$ is the child’s W score on each of the WJ literacy subtests, dfc_{ij} is the running variable that denotes days from cutoff based on the child’s DOB, Tx_{ij}^{exp} is a dummy variable where a value of 1 indicates that the child made the cutoff for school enrollment based on the child’s DOB, $dfc_{ij} \times Tx_{ij}^{exp}$ is an interaction term that accounts for any change in the slope of the regression line for children who made the cutoff relative to those who missed the cutoff, $covariates_{ij}$ is a vector of covariates that includes baseline characteristics, sfe_j are school fixed effects, ε_{ij} is an error term, subscript i denotes each child, and subscript j denotes each school.

TABLE 5

ITT Estimates of Schooling Effects on Woodcock Johnson Literacy Subtests

Variables	Pre-K effect				Kindergarten (K) effect				1st grade effect			
	(1a)	(1b)	(1c)	(1d)	(2a)	(2b)	(2c)	(2d)	(3a)	(3b)	(3c)	(3d)
	LW	SA	PV	PC	LW	SA	PV	PC	LW	SA	PV	PC
Linear	18.37*	6.10	1.09	7.71	30.84***	4.62	4.66	14.81*	28.14**	4.02	-0.65	21.28**
	(7.63)	(4.51)	(3.42)	(4.96)	(8.82)	(4.35)	(3.08)	(6.43)	(9.81)	(3.44)	(3.05)	(6.94)
<i>Effect size</i>	<i>0.29</i>	<i>0.17</i>	<i>0.05</i>	<i>0.20</i>	<i>0.44</i>	<i>0.14</i>	<i>0.20</i>	<i>0.30</i>	<i>0.39</i>	<i>0.15</i>	<i>-0.03</i>	<i>0.39</i>
Local linear	15.72*	9.41 ⁺	2.66	3.51	26.70*	15.37*	6.65 ⁺	9.45	21.11 ⁺	1.96	-1.22	21.38**
	(7.36)	(5.31)	(3.74)	(5.93)	(11.42)	(6.49)	(3.73)	(8.40)	(11.47)	(3.93)	(3.26)	(7.86)
Quadratic	5.49	5.96	5.52	0.84	13.04	8.53	3.39	3.19	15.46	2.24	-2.48	16.62 ⁺
	(9.40)	(6.60)	(4.22)	(6.27)	(13.19)	(6.46)	(4.40)	(9.45)	(14.20)	(5.10)	(4.48)	(9.32)
<i>N</i> (linear, quadratic)	226	226	226	225	220	220	220	220	203	203	203	203
<i>N</i> (local linear)	189	162	136	127	145	116	137	145	145	177	142	141
Optimal bandwidth (days; local linear)	149	128	103	98	114	90	108	115	123	157	121	121

Robust standard errors are in parentheses. Outcomes are measured in W score units on an equal-interval scale.

LW = letter-word identification; SA = sound awareness; PV = picture vocabulary; PC = passage comprehension.

Bandwidth is 180 days for linear and quadratic estimates. Optimal local linear bandwidth was generated using the method developed by Calonico et al. (2014). Number of observations is a function of bandwidth selection differences between linear and local linear specifications.

⁺ $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$.

Presenting results from multiple estimation methods can provide valuable information indicating the stability and consistency of our estimates. To that end, we also present TOT estimates using a method called two-stage least squares (see McEwan & Shapiro, 2008, for an example), which uses the child's expected treatment status (based on the child's DOB) as an instrumental variable to predict the child's actual treatment status. Specifically, we estimate the following regression:

$$\begin{aligned}
 Tx_{ij}^{act} &= \alpha_0 + \alpha_1 Tx_{ij}^{exp} + \alpha_2 dfc_{ij} + \alpha_3 dfc_{ij} \\
 &\quad \times Tx_{ij}^{exp} + \alpha_4 covariates_{ij} + \alpha_5 sfe_j + \varepsilon_{ij} \\
 outcome_{ij} &= \beta_0 + \beta_1 \widehat{Tx}_{ij}^{act} + \beta_2 dfc_{ij} + \beta_3 dfc_{ij} \\
 &\quad \times Tx_{ij}^{exp} + \beta_4 covariates_{ij} + \beta_5 sfe_j + \varepsilon_{ij}
 \end{aligned}$$

where Tx_{ij}^{act} is the child's actual treatment status where a value of 1 indicates that the child made the cutoff for school enrollment, Tx_{ij}^{exp} is the child's expected treatment status, dfc_{ij} is the quantitative assignment variable that denotes days from cutoff based on the child's DOB, $dfc_{ij} \times Tx_{ij}^{exp}$ is an interaction term that accounts for any change in the slope of the regression line for children who made the cutoff relative to those who missed the cutoff, $covariates_{ij}$ is a vector of covariates that includes baseline characteristics, sfe_j are school fixed effects, ε_{ij} is an error term, subscript i denotes each child, and subscript j denotes each school. In the second stage, $outcome_{ij}$ is the child's W score on each of the WJ literacy subtests, and \widehat{Tx}_{ij}^{act} is the predicted value of Tx_{ij}^{act} that was generated from the first stage.

Results

Intent-To-Treat Estimates

Reduced-form ITT estimates of schooling on early literacy are presented in Table 5. We provided estimates from linear, local linear, and quadratic specifications. The quadratic and local linear point estimates were within the 95% confidence intervals for the associated linear estimates, suggesting that the estimates were robust to RD specification and were not significantly different from each other. That said, nonlinear, higher-order polynomial specifications do not appear to model our data well, likely due to the substantial noise associated with our relatively small sample; a linear specification is therefore the preferred approach. Data indicated that the link between schooling and literacy depended on grade level and literacy subskill. Consistent with our original prediction, schooling had moderate positive effects on decoding at all three grade levels (effect sizes: pre-K = .29; kindergarten = .44; first grade = .39). However, differences in the magnitude of the schooling effect on decoding as a function of grade level were not statistically significant. Figure 2 presents a graphical representation of the linear effect of schooling on decoding at each grade level. Partially confirming our hypothesis, kindergarten and first grade schooling predicted comprehension skills (effect sizes: kindergarten = .30, first grade = .39). Differences in the magnitude of the schooling effect on comprehension as a function of grade level were not statistically significant. Schooling was unrelated to expressive vocabulary at any grade level, consistent with our original prediction. Contrary

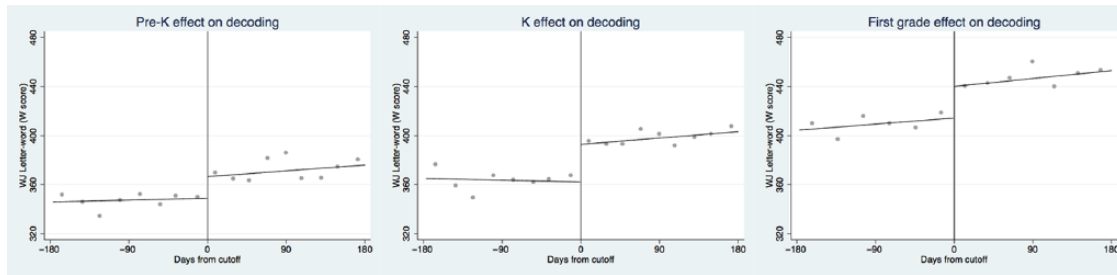


FIGURE 2. *Schooling effects (intent-to-treat) on decoding skills.*

Number of bins were automatically selected to mimic the variation in the data using the method developed by Calonico, Cattaneo, and Titiunik (2015). Dots indicate the sample average of the Woodcock Johnson Letter-Word Identification W score within each bin. Solid line represents a linear regression fit.

to our hypothesis, schooling was unrelated to sound awareness at any grade level.

Treatment on the Treated Estimates

TOT estimates of schooling on early literacy are presented in Table 6. While noncompliance rates were between 4.8% and 5.1% for the entire sample, noncompliance occurs with greater frequency near the cutoff, which is not surprising as treatment manipulation would be expected to occur at or near the cutoff. Accordingly, the TOT estimates were approximately twice as large as the corresponding ITT estimates, and the pattern of significance was identical to that of the ITT estimates. The large F -statistics indicate that the first-stage regression is strong; that is, the child's expected treatment status is a relevant instrument for the child's actual treatment status. As before, the quadratic and local linear point estimates were within the 95% confidence intervals for the associated linear estimates. While effect sizes cannot be computed in a two-stage least squares regression model, ITT effect sizes can be used as a reference point for approximating the corresponding effect sizes for the TOT estimates.

Robustness Checks

In addition to presenting estimates from ITT and TOT estimation methods, additional robustness checks were conducted in order to assess the degree to which our estimates were sensitive to different methods of RD implementation. Given the relatively small sample size in the current investigation, these sensitivity analyses are particularly important for demonstrating whether the estimates are real or simply artifacts due to noise. These robustness checks are presented in the Online Supplemental Material. In Figure S1, we demonstrate that the 95% confidence intervals associated with each bandwidth—across all grade levels and literacy skills—overlap with each other, indicating that our estimates are robust to bandwidth selection. In Table S2, we demonstrate that the point estimates and pattern of significance are virtually identical to the main analysis when excluding the small number of children enrolled in schools that only appear on

one side of the cutoff. In Table S3, we address potential concerns regarding attrition by demonstrating that the results are very similar when examining a restricted sample of children for which data are available at all three time points (pre-K, kindergarten, and first grade). Finally, in Table S4, we demonstrate that the pattern of significance does not change when we use the standard score instead of the W score as the outcome metric.

Discussion

The current study used a regression discontinuity design to examine the nature and magnitude of the effect of schooling on a range of early literacy skills during the transition to school period. Findings were broadly consistent with earlier school cutoff and RD investigations of early literacy and were robust to a series of sensitivity tests. Notably, the longitudinal data used in the present study allowed us to generate new insights into the specificity of the link between schooling and literacy as a function of both grade level and literacy subskill. Results revealed moderate positive effects of prekindergarten, kindergarten, and first grade schooling on decoding, while kindergarten and first grade schooling predicted comprehension skills. There was no significant effect of schooling at any grade level on expressive vocabulary and sound awareness.

Schooling Effects on Decoding and Comprehension

Consistent with our original prediction, our results indicate that pre-K, kindergarten, and first grade schooling each have moderate positive effects on children's decoding skills, over and above age-related change. Our results are consistent with earlier findings that demonstrate schooling effects on decoding skills (Burrage et al., 2008; Christian et al., 2000; Gormley et al., 2005; Weiland & Yoshikawa, 2013; Wilson et al., 2013). It is important to note that the treatment was not a special intervention or enrichment program. Rather, the treatment was a research-based core literacy curriculum implemented in public schools. The results demonstrate that the early learning experiences that children

TABLE 6
TOT (Treatment on the Treated) Estimates of Schooling Effects on Woodcock-Johnson Literacy Subtests

Variables	Pre-K effect						Kindergarten (K) effect						1st grade effect											
	(1a)		(1b)		(1c)		(1d)		(2a)		(2b)		(2c)		(2d)		(3a)		(3b)		(3c)		(3d)	
	LW	SA	LW	SA	PV	PC	LW	PC	LW	SA	LW	SA	PV	PC	LW	SA	LW	SA	PV	SA	PV	SA	PC	
Linear	34.83*** (13.41)	11.57 (7.80)	2.06 (6.21)	14.61 (8.96)	59.96*** (15.09)	8.98 (7.66)	9.06 (5.86)	28.79** (11.56)	51.72*** (14.74)	7.39 (5.70)	-1.20 (5.28)	39.12*** (10.83)												
Local linear	24.88* (11.11)	13.07 ⁺ (6.92)	1.59 (5.11)	5.09 (8.30)	40.22* (17.70)	14.36 ⁺ (7.49)	10.44 ⁺ (6.01)	14.22 (13.11)	42.43** (15.65)	3.48 (7.03)	-2.92 (6.00)	38.75** (11.68)												
Quadratic	10.35 (16.31)	11.23 (11.02)	10.40 (7.97)	1.58 (11.16)	25.56 (22.91)	16.71 (11.01)	6.65 (8.33)	6.25 (17.25)	30.31 (23.28)	4.40 (9.29)	-4.87 (8.21)	32.60** (15.23)												
F-statistic (linear)	F(1,225)= 36.74***	F(1,225)= 36.74***	F(1,225)= 36.74***	F(1,224)= 36.76***	F(1,219)= 33.59***	F(1,219)= 33.59***	F(1,219)= 33.59***	F(1,219)= 33.59***	F(1,202)= 33.56***	F(1,202)= 33.56***	F(1,202)= 33.56***	F(1,202)= 33.56***												
N (linear, quadratic)	226	226	226	225	220	220	220	220	203	203	203	203												
N (local linear)	204	216	204	136	131	203	135	140	202	202	162	202												

Robust standard errors are in parentheses. Outcomes are measured in W score units on an equal-interval scale.

LW = letter-word identification; SA = sound awareness; PV = picture vocabulary; PC = passage comprehension.

Bandwidth is 180 days for linear and quadratic estimates. Optimal bandwidth is used for local linear estimates. Number of observations is a function of bandwidth selection differences between linear and local linear specifications.

⁺ $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$.

typically receive in formal public school environments can significantly promote beginning literacy skills.

Our observed pre-K effect on decoding skills (.29 SD for ITT) is smaller than those reported in previous studies using the same decoding measure (.62 SD in Weiland & Yoshikawa, 2013; .79 SD in Gormley et al., 2005). One explanation for this difference may be due to the fact that schooling effects may be attenuated in our racially homogeneous and upper-middle SES sample, compared to the more diverse samples studied in previous empirical investigations. Indeed, because our sample was drawn from relatively upper-middle SES backgrounds, the external validity of our findings is a limitation of the present study. That said, our observed pattern of positive schooling effects on decoding is identical to that of previous school cutoff investigations using samples with a similar sociodemographic profile (e.g., Burrage et al., 2008; Skibbe et al., 2011), indicating an important degree of replicability across analytical techniques. Nevertheless, previous RD investigations on public pre-K and Head Start programs have provided important insights into the effects of schooling on socioeconomically diverse populations, and we believe it is important to continue to explore how schooling impacts those children who might benefit from formal instruction the most.

The present study also adds to the literature on early literacy development by demonstrating that children's early comprehension skills were predicted by kindergarten and first grade schooling but not pre-K. Previous research using a subset of classroom observation data from the same study indicates that pre-K and kindergarten children received very similar amounts of instruction in decoding and comprehension (Skibbe, Hindman, Connor, Housey, & Morrison, 2013). The fact that we observed a kindergarten effect but not a pre-K effect on comprehension, despite similar amounts of instruction in both grades, might be due to differences in the nature or quality of kindergarten instruction. Another possibility is that older children are more cognitively mature and thus better able to acquire comprehension skills compared to younger children, regardless of the nature or quality of instruction (Del Giudice, 2014).

Because we examined schooling effects at each of three different grade levels on the same set of literacy subskills, an important question is whether children's performance on a given literacy skill depends on the particular grade level of schooling. For example, because decoding is an important prerequisite skill for fluent comprehension, we might expect decoding skills to be more strongly emphasized in the earlier grades (pre-K and kindergarten), and comprehension to be shaped more strongly by instruction in first grade. However, we found that the magnitude of effects across grade levels were not significantly different from each other. For example, the effects of first grade on comprehension were not statistically larger than the effects of kindergarten on comprehension.

There are several possible explanations as to why we did not observe any differences in the magnitude of the schooling effect as a function of grade level. It is possible that the relatively small sample size did not provide adequate power in order to detect such differences. However, it is also possible that there really are no meaningful differences as a function of grade level within a particular literacy subskill, at least in the early elementary grades examined in the present study. Such an interpretation would be contrary to our prediction that the effects of schooling should match the developmental trajectories of each literacy subskill (e.g., larger effects of comprehension in later grades, larger effects of decoding in earlier grades, etc.), but it might indicate that early schooling—regardless of grade level—exerts similar positive effects on children's performance on discrete dimensions of early literacy. Future research should seek to leverage additional sources of classroom- and school-level data that were not available in the present investigation—such as amounts and types of literacy instruction as well as other measures of classroom functioning—in order to assess the validity of each interpretation.

Null Effects on Expressive Vocabulary and Sound Awareness

The null effects of schooling on expressive vocabulary were consistent with our hypothesis and confirm previous findings demonstrating that age-related change is a stronger predictor of expressive vocabulary growth than schooling (Skibbe et al., 2011). Notably, the standard errors associated with our vocabulary estimates are smaller than those for decoding, sound awareness, and comprehension across all three grade levels and both estimation methods, indicating that these null effects are likely to be real and not simply an artifact of noise. That said, it is important to note that previous research using RD has found a significant pre-K effect on *receptive* vocabulary (Weiland & Yoshikawa, 2013). Taken together, these results seem to indicate that the effects of schooling might depend on the type of vocabulary skill being taught and assessed. Alternatively, the difference between our results and findings from previous studies might reflect the differential emphasis between school districts in promoting different aspects of vocabulary during the early grades.

The null effects we observed on sound awareness were, however, not consistent with our original prediction and contrast with previous school cutoff studies that demonstrated a positive effect of first grade on phonemic awareness (e.g., Christian et al., 2000; Morrison et al., 1995). An alternate model specification that excluded crossover children yielded results indicating that pre-K and kindergarten schooling each had unique positive effects on sound awareness skills. However, because this result was not robust to the ITT and TOT specifications more commonly implemented in RD

designs, it is difficult to regard these effects as meaningful in our sample.

Limitations and Opportunities for Future Research

A general limitation of the present study was the lack of access to observational data for all classrooms in which students received instruction. Because data collection occurred during the initial years of Reading First and prior to the implementation of the Common Core State Standards, observational data would have allowed us to characterize more precisely the nature of literacy instruction in the schools and classrooms in our sample in order to facilitate comparisons with other studies. Individual differences in teacher expertise in delivering effective literacy instruction might also have affected the results. In addition, data on how teachers actually delivered literacy instruction (i.e., fidelity of implementation), as well as the extent to which teachers provided opportunities for children to practice literacy skills, would help us better understand our findings. Children receiving instruction that is primarily code-focused (targeting decoding and sound awareness) experience a different pattern of literacy outcomes compared to children receiving meaning-focused instruction (targeting comprehension and vocabulary) (e.g., Connor, Morrison, & Katch, 2004; Connor et al., 2006). While we somewhat mitigated this limitation by controlling for school fixed effects in all of our estimated models, a more nuanced exploration of the specific factors that might be underlying our observed schooling effects would include an examination of classroom- and teacher-level data as well as data on implementation fidelity.

The technical advances in RD in recent years have been accompanied by a greater awareness of threats to internal validity in the specific context of early schooling that could affect the precision and interpretation of RD estimates (e.g., Lipsey et al., 2015). One potential source of bias is that the children who agreed to participate in the study may be qualitatively different from children who did not participate. Therefore, any schooling effect that we observe may understate or overstate the true effect, depending on the nature of the difference (if any) between the two groups. Similarly, differential attrition due to children who were enrolled in school one year but left the sample prior to the next year might also lead to biased estimates of schooling effects. Table S3 demonstrates that the effects are very similar when examining a restricted sample of students for whom data were available at all time points, and frequent interactions with teachers and children led the study team to conclude that participating children did not significantly differ from non-participating children (Connor et al., 2006). Nevertheless, a rigorous analysis of administrative data would be the preferred approach for determining more precisely the effects, if any, of differential attrition.

The study from which the data were drawn was not originally designed as a quasi-experimental investigation of schooling. A retrospective power analysis indicated that our sample size was sufficient to detect the observed effects on our decoding and comprehension measures, but might not have been sufficient to detect more nuanced effects. However, by adopting the more conservative ITT approach as well as focusing on just those estimates that were robust to the largest number of sensitivity analyses, we believe that we have highlighted the schooling effects that are most likely to be meaningful in our small-scale RD implementation. Future investigations should seek to replicate these findings using a larger and more racially and socioeconomically diverse sample of children in order to confirm or extend our observations of the differential impact of schooling on literacy development during the school transition period.

Conclusion

The current study used a regression discontinuity design to examine the nature and magnitude of schooling effects on a range of literacy skills during the transition to school. Results indicated moderate positive effects of prekindergarten, kindergarten, and first grade schooling on decoding, while kindergarten and first grade schooling predicted comprehension skills. There was no significant effect of schooling at any grade level on expressive vocabulary and sound awareness. These findings highlight the heterogeneity of schooling effects on early literacy skills. Future investigation of the nature, quality, and quantity of literacy instruction would provide insights into the precise mechanisms of these schooling effects, with potential implications for designing and implementing literacy instruction during the school transition period.

Acknowledgments

We thank Caroline Weber for her consultation on the analytic approach used in the study. We would also like to thank the families and children who participated in our longitudinal study of literacy development from which the data were collected.

Funding

This research was supported in part by a grant from the National Institute of Child Health and Human Development (HD27176-R21) to the second author.

ORCID iD

Matthew H. Kim  <https://orcid.org/0000-0002-1436-3596>

References

- Bassok, D., Latham, S., & Rorem, A. (2016). Is kindergarten the new first grade? *AERA Open*, 1(4), 1–31. doi: 10.1177/2332858415616358

- Bassok, D., & Reardon, S. F. (2013). "Academic redshirting" in kindergarten: Prevalence, patterns, and implications. *Educational Evaluation and Policy Analysis*, 35(3), 283–297. doi: 10.3102/0162373713482764
- Bloom, H. (2009). *Modern regression discontinuity analysis* (Working Paper). New York: MDRC.
- Borman, G. D., Slavin, R. E., Cheung, A. C. K., Chamberlain, A. M., Madden, N. A., & Chambers, B. (2007). Final reading outcomes of the national randomized field trial of Success for All. *American Educational Research Journal*, 44(3), 701–731. doi: 10.3102/0002831207306743
- Burrage, M. S., Ponitz, C. C., McCreedy, E. A., Shah, P., Sims, B. C., Jewkes, A. M., & Morrison, F. J. (2008). Age- and schooling-related effects on executive functions in young children: A natural experiment. *Child Neuropsychology*, 14, 510–524. doi: 10.1080/09297040701756917
- Cahan, S., & Cohen, N. (1989). Age versus schooling effects on intelligence development. *Child Development*, 60, 1239–1249. doi: 10.2307/1130797
- Calonico, S., Cattaneo, M. D., Farrell, M. H., & Titiunik, R. (2016). *Regression discontinuity designs using covariates* (Working Paper). Ann Arbor, MI: University of Michigan.
- Calonico, S., Cattaneo, M. D., & Titiunik, R. (2014). Robust non-parametric confidence intervals for regression-discontinuity designs. *Econometrica*, 82(6), 2295–2326. doi: 10.3982/ECTA11757
- Calonico, S., Cattaneo, M. D., & Titiunik, R. (2015). Optimal data-driven regression discontinuity plots. *Journal of the American Statistical Association*, 110, 1753–1769. doi: 10.1080/01621459.2015.1017578
- Christian, K., Morrison, F. J., Frazier, J. A., & Massetti, G. (2000). Specificity in the nature and timing of cognitive growth in kindergarten and first grade. *Journal of Cognition and Development*, 1(4), 429–448. doi: 10.1207/S15327647JCD0104_04
- Connor, C. M. (2016). A lattice model of the development of reading comprehension. *Child Development Perspectives*, 10(4), 269–274. doi: 10.1111/cdep.12200
- Connor, C. M., Morrison, F. J., & Katch, L. E. (2004). Beyond the reading wars: Exploring the effect of child-instruction interactions on growth in early reading. *Scientific Studies of Reading*, 8(4), 305–336. doi: 10.1207/s1532799xssr0804_1
- Connor, C. M., Morrison, F. J., & Slominski, L. (2006). Preschool instruction and children's emergency literacy growth. *Journal of Educational Psychology*, 98(4), 665–689. doi: 10.1037/0022-0663.98.4.665
- Connor, C. M., Son, S-H., Hindman, A. H., & Morrison, F. J. (2005). Teacher qualifications, classroom practices, family characteristics, and preschool experience: Complex effects on first graders' vocabulary and early reading outcomes. *Journal of School Psychology*, 43, 343–375. doi: 10.1016/j.jsp.2005.06.001
- Del Giudice, M. (2014). Middle childhood: An evolutionary-developmental synthesis. *Child Development Perspectives*, 8(4), 193–200. doi: 10.1111/cdep.12084
- Dickinson, D. K., McCabe, A., Anastasopoulos, L., Peisner-Feinberg, E. S., & Poe, M. D. (2003). The comprehensive language approach to early literacy: The interrelationships among vocabulary, phonological sensitivity, and print knowledge among preschool-aged children. *Journal of Educational Psychology*, 95(3), 465–481. doi: 10.1037/0022-0663.95.3.465
- Farver, J. A. M., Lonigan, C. J., & Eppe, S. (2009). Effective early literacy skill development for young Spanish-speaking English language learners: An experimental study of two methods. *Child Development*, 80(3), 703–719. doi: 10.1111/j.1467-8624.2009.01292.x
- Gelman, A., & Imbens, G. (2014). *Why high-order polynomials should not be used in regression discontinuity designs* (NBER Working Paper No. 20405). Cambridge, MA: National Bureau of Economic Research.
- Gormley, W. T., Gayer, T., Phillips, D., & Dawson, B. (2005). The effects of universal pre-K on cognitive development. *Developmental Psychology*, 41(6), 872–884. doi: 10.1037/0012-1649.41.6.872
- Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore, MD: Brookes.
- Imbens, G. W., & Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics*, 142, 615–635. doi: 10.1016/j.jeconom.2007.05.001
- Jacob, R., Zhu, P., Somers, M., & Bloom, H. (2012). *A practical guide to regression discontinuity* (Working Paper). New York: MDRC.
- Jaffe, L. E. (2009). *Development, interpretation, and application of the W score and the relative proficiency index* (Woodcock-Johnson III Assessment Service Bulletin No. 11). Rolling Meadows, IL: Riverside Publishing.
- Juel, C., Biancarosa, G., Coker, D., & Deffes, R. (2003). Walking with Rosie: A cautionary tale of early reading instruction. *Educational Leadership*, 60(7), 12–18.
- Keenan, J. M., Betjemann, R. S., & Olson, R. K. (2008). Reading comprehension tests vary in the skills they assess: Differential dependence on decoding and oral comprehension. *Scientific Studies of Reading*, 12(3), 281–300. doi: 10.1080/10888430802132279
- Kendeou, P., van den Broek, P., White, M. J., & Lynch, J. S. (2009). Predicting reading comprehension in early elementary school: The independent contributions of oral language and decoding skills. *Journal of Educational Psychology*, 101(4), 765–778. doi: 10.1037/a0015956
- Kim, M. H., Bell, L. H., & Morrison, F. J. (2011, March). *Math in kindergarten classrooms: Effects of children's executive functioning and instruction on early math achievement*. Presented at the biennial meeting of the Society for Research in Child Development, Montreal, Canada.
- Lincove, J. A., & Painter, G. (2006). Does the age that children start kindergarten matter? Evidence of long-term educational and social outcomes. *Educational Evaluation and Policy Analysis*, 28(2), 153–179. doi: 10.3102/01623737028002153
- Lipsey, M. W., Weiland, C., Yoshikawa, H., Wilson, S. J., & Hofer, K. G. (2015). The prekindergarten age-cutoff regression-discontinuity design: Methodological issues and implications for application. *Educational Evaluation and Policy Analysis* 37(3), 296–313. doi: 10.3102/01623737028002153
- Lonigan, C. J., Burgess, S. R., & Anthony, J. L. (2000). Development of emergent literacy and early reading skills in preschool children: Evidence from a latent-variable longitudinal study. *Developmental Psychology*, 36(5), 596–613. doi: 10.1037/0012-1649.36.5.596
- McEwan, P. J., & Shapiro, J. S. (2008). The benefits of delayed primary school enrollment: Discontinuity estimates using exact

- birth dates. *The Journal of Human Resources*, 43(1), 1–29. doi: 10.3368/jhr.43.1.1
- Morrison, F. J., Smith, L., & Dow-Ehrensberger, M. (1995). Education and cognitive development: A natural experiment. *Developmental Psychology*, 31(5), 789–799. doi: 10.1037/0012-1649.31.5.789
- National Early Literacy Panel (NELP) (2008). *Developing early literacy: Report of the National Early Literacy Panel*. Washington, DC: National Institute for Literacy.
- Neuman, S. B., & Dwyer, J. (2009). Missing in action: Vocabulary instruction in pre-K. *The Reading Teacher*, 62(5), 384–392. doi: 10.1598/RT.62.5.2
- Phillips, B. M., Clancy-Menchetti, J., & Lonigan, C. J. (2008). Successful phonological awareness instruction with preschool children: Lessons from the classroom. *Topics in Early Childhood Special Education*, 28(1), 3–17. doi: 10.1177/0271121407313813
- Phillips, D. A., Gormley, W. T., & Lowenstein, A. E. (2009). Inside the prekindergarten door: Classroom climate and instructional time allocation in Tulsa's pre-K programs. *Early Childhood Research Quarterly*, 24, 213–228. doi: 10.1016/j.ecresq.2009.05.002
- Roberts, J., Jurgens, J., & Burchinal, M. (2005). The role of home literacy practices in preschool children's language and emergent literacy skills. *Journal of Speech, Language, and Hearing Research*, 48, 345–359. doi: 10.1044/1092-4388(2005/024)
- Schwartz, R. M. (2005). Literacy learning of at-risk first-grade students in the Reading Recovery early intervention. *Journal of Educational Psychology*, 97(2), 257–267. doi: 10.1037/0022-0663.97.2.257
- Sénéchal, M. (1997). The differential effect of storybook reading on preschoolers' acquisition of expressive and receptive vocabulary. *Journal of Child Language*, 24, 123–138.
- Skibbe, L. E., Connor, C. M., Morrison, F. J., & Jewkes, A. M. (2011). Schooling effects on preschoolers' self-regulation, early literacy, and language growth. *Early Childhood Research Quarterly*, 26, 42–49. doi: 10.1016/j.ecresq.2010.05.001
- Skibbe, L. E., Grimm, K. J., Bowles, R. P., & Morrison, F. J. (2012). Literacy growth in the academic year versus summer from preschool through second grade: Differential effects of schooling across four skills. *Scientific Studies of Reading*, 16(2), 141–165. doi: 10.1080/10888438.2010.543446
- Skibbe, L. E., Hindman, A. H., Connor, C. M., Housey, M., & Morrison, F. J. (2013). Relative contributions of prekindergarten and kindergarten to children's literacy and mathematics skills. *Early Education and Development*, 24(5), 687–703. doi: 10.1080/10409289.2012.712888
- Skibbe, L. E., Montroy, J. J., Bowles, R. P., & Morrison, F. J. (2018). Self-regulation and the development of literacy and language achievement from preschool through second grade. *Early Childhood Research Quarterly*. Advance online publication. doi: 10.1016/j.ecresq.2018.02.005
- Varnhagen, C. K., Morrison, F. J., & Everall, R. (1994). Age and schooling effects in story recall and story production. *Developmental Psychology*, 30(6), 969–979. doi: 10.1037/0012-1649.30.6.969
- Watts, T. W., Spanier, M., & Duncan, G. J. (2014, April). *Predicting adolescent math achievement with preschool math skills*. Paper presented at the annual meeting of the American Educational Research Association, Philadelphia, PA.
- Weiland, C., & Yoshikawa, H. (2013). Impacts of a prekindergarten program on children's mathematics, language, literacy, executive function, and emotional skills. *Child Development*, 84(6), 2112–2130. doi: 10.1111/cdev.12099
- Whitehurst, G. J., & Lonigan, C. J. (1998). Child development and emergent literacy. *Child Development*, 69(3), 848–872. doi: 10.1111/j.1467-8624.1998.tb06247.x
- Wilson, S. J., Dickinson, D. K., & Rowe, D. W. (2013). Impact of an Early Reading First program on the language and literacy achievement of children from diverse language backgrounds. *Early Childhood Research Quarterly*, 28, 578–592. doi: 10.1016/j.ecresq.2013.03.006
- Wong, V. C., Cook, T. D., Barnett, W. S., & Jung, K. (2008). An effectiveness-based evaluation of five state prekindergarten programs. *Journal of Policy Analysis and Management*, 27(1), 122–154. doi: 10.1002/pam.20310
- Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III tests of achievement*. Itasca, IL: Riverside.

Authors

MATTHEW H. KIM is a research scientist at the Institute for Learning and Brain Sciences and a teaching associate in the College of Education at the University of Washington. His research examines the nature and development of executive functions and motivation during the school transition period using behavioral, ERP, and causal inference methods, and how these cognitive processes relate to early academic success.

FREDERICK J. MORRISON is a professor of psychology at the University of Michigan. His current research focuses on the nature and sources of literacy acquisition and executive functions in children during the transition to school.