ELIZABETH COYNE CROWE, Florida Center for Reading Research, CAROL MCDONALD CONNOR, Florida State University, and YAACOV PETSCHER, Florida State University

Examining the Core: Relations Among Reading Curricula, Poverty, and First-Through Third-Grade Reading Achievement

Abstract: Policy changes at the federal and state level are endeavoring to improve student achievement at schools serving children from lower-SES homes. One important strategy is the focus on using evidence-based core reading curricula to provide a consistent framework for instruction across schools. However, rarely have these curricula undergone rigorous comparative testing. Therefore, the purpose of this study was to compare the effects of six core reading curricula on oral reading fluency growth, while appraising whether these effects differ by grade level and for children living in lower socioeconomic (SES) households. Over 30,000 students in first- through third-grade Florida Reading First classrooms comprise this academically and economically diverse cross-sectional. Hierarchical Linear Modeling was used to model latent growth curves for students' reading fluency scores over the school year. Growth curves revealed differences across curricula as well as between students of lower and higher SES, suggesting that reading fluency growth trajectories for curricula varied depending on student SES and grade level. Findings indicate that while there are similarities among curricula, they sometimes differ in their ability to promote reading skill growth. Differences by grade level and SES were also detected. However, many of these differences were small. Implications for the use of curriculum as a conduit for improving reading instruction are discussed.

Socioeconomic status (SES) has been identified as a unique contributor to academic achievement (Duncan & Brooks-Gunn, 2000; Duncan & Raudenbush, 1999; Duncan, Yeung, Brooks-Gunn, & Smith, 1998; Hart & Risley, 1995; Raudenbush, 2004). On national literacy assessments, students in lower SES homes continue to score lower than students in homes that do not qualify for free lunch programs (Lee, Griggs, & Donahue, 2007). Students' SES when entering school not only influences their early academic outcomes, but also the sustainability of average and above-average achievement (Duncan et al., 1998; Wyner, Bridgeland, & Diiulio, 2007). Specifically, children from lower-SES homes often begin school with weaker language and literacy skills than do children from higher-SES homes (Entwisle & Alexander, 1993; Hart & Risley, 1995), and students who are not at grade level upon completion of first grade have dramatically lower chances of being on or above grade level later in elementary school (Spira, Bracken, & Fischel, 2005; Wyner et al., 2007). Taken together, these factors may contribute to the high incidence of failure and

Journal of Direct Instruction, Vol. 10, pp. 49-75. Address correspondence to Elizabeth Crowe at ecrowe@fcrr.org.

Reprinted with permission of Elsevier Ltd. Original citation: Crowe, E. C., Connor, C. M., and Petscher, Y. Examining the core: Relations among reading curricula, poverty, and first through third grade reading achievement. *Journal of School Psychology, 47* (2009), 187–214.

delayed reading skill acquisition among children living in lower-SES homes.

Research on academic difficulties related to socioeconomic factors points to SES influences before and after entrance to school (Evans, 2004; Kozol, 1991; Lee & Burkam, 2002; Rothstein, 2004). Prior to school, students from lower-SES families tend to have fewer literacy opportunities compared to their higher-SES peers. Such differences may include having fewer books in the home (Evans, 2004; Lee & Burkam, 2002; Vernon-Feagans, Hammer, Miccio, & Manlove, 2002; Whitehurst & Lonigan, 1998) and attending lower-quality preschools, if students have these opportunities at all (McCoach, O'Connoll, Reis, & Levitt, 2006; NICHD-ECCRN, 2002). Such differences may be related to the acquisition of early reading skills (Arnold & Doctoroff, 2003; Duncan, Yeung, Brooks-Gunn, & Smith, 1998; Lee & Burkam, 2002; Yeung, Linver, & Brooks-Gunn, 2002), and may also lead to lower achievement in subsequent grades if efforts are not made to accelerate students' learning (Wyner et al., 2007).

Once students begin school, differences between lower- and higher-SES students may be as much related to poor-quality academic experiences and interactions as to home characteristics (Aikens & Barbarin, 2008). Specific lags associated with lower-SES students include a delayed ability to identify letters and words and a lack of phonological sensitivity (Whitehurst & Lonigan, 1998) often thought to be attributable to less experienced teachers (Haycock, 2000), less family involvement in school and classroom activities (Evans, 2004), and attendance at schools with fewer resources resulting in fewer opportunities to develop language and literacy skills (Connor, Son, Hindman, & Morrison, 2005).

Together, both the home and school represent important sources of influence, which may be responsible for SES differences in academic outcomes (Aikens & Barbarin, 2008; Evans, 2004; Lee & Burkam, 2002). Modifying the influence of SES on social and academic outcomes has proven to be complex (Katz, Kling, & Liebman 2003; Ladd and Ludwig 2003). Therefore, raising the achievement of lower-SES students, at least on a short-term scale, may be best approached through the alteration of classroom and school instructional practices over which educators have direct control.

Reforms of classroom reading instruction for lower-SES students indicate that instructional strategies can be designed to improve literacy skills (Pressley et al., 2001; Pressley, Wharton-McDonald, Mistretta-Hapston, & Echevarria, 1998; Taylor, Pearson, Clark, & Walpole, 2000; Wharton-McDonald, Pressley, & Hampston, 1998). In these studies, lower-SES schools were able to overcome low achievement, with students making achievement gains commensurate with national norms. Strategies utilized included small, homogeneous, teacher-managed, skill-based reading groups (Pressley et al., 1998; Pressley et al., 2001; Taylor et al., 2000; Wharton-McDonald et al., 1998), as well as instruction adapted and refined based on frequent assessments of student skills (Connor, Morrison, Fishman, Schatschneider, & Underwood, 2007). These results suggest that, although the progress of students lower-SES is often poor (Juel, 1988), growth trajectories for such students may be improved through appropriate instruction (Foorman et al., 2006; Foorman & Torgesen, 2001).

One example of a large-scale response to the need to reform reading instruction is the Reading First initiative, enacted to promote overall reading achievement while specifically targeting schools that displayed a lack of progress in reading. This policy-based program introduced a "no excuses" perspective on student literacy outcomes and scientifically based standards of instruction through an emphasis on key reading components as outlined by the Reading First implementation report (USDOE, 2006). In an effort to ensure consistently high-quality reading instruction,

a central requirement of participation in the Reading First program was the adoption of scientifically based curricula. While specific curricula were not mandated, the choice of a reading program that adequately addressed the five essential components of reading instruction outlined by the National Reading Panel's report (phonics, phonemic awareness, fluency, comprehension, and vocabulary; NRP, 2000) was required and, in Florida, implementation was monitored through randomly selected site visits by Reading First teams. Although efforts to improve reading instruction were coupled with professional development and teacher training, considerable responsibility was placed on the curriculum to improve the quality of instruction.

One complication to the use of curriculum as a component of reading reform lies in the possibility that the effectiveness of any one curriculum may depend on many factors, the most prominent of which may be characteristics of the students themselves (e.g., skills, background, and SES). These distinctions may interact with the instruction they receive, resulting in differentiated patterns of reading-skill growth across children. Therefore, reading curricula may be more or less effective for students with specific child-level characteristics (e.g., lower-SES), depending on the skill emphases and instructional plan of a particular curriculum.

Published core curricula, although based on scientific research, likely target typically developing children. Many children living in lower-SES homes begin school with weaker language and literacy skills than their more affluent peers (McCoach et al., 2006; NICHDECCRN, 2002; USDOE, 2006). Hence, curricula targeted toward typically developing, middle- and higher-SES children may not provide the generally more intensive and targeted instruction needed by children with weaker initial skills (Connor, Piasta, Glasney, Schatschneider, Fishman, Underwood, 2007; Torgesen et al., 2001). Therefore, research clarifying the effects of different reading curricula on students of varying SES is of vital importance to assuring an equitable educational experience. The following research questions guided the current investigation: 1) What are the effects of different core curricula on children's reading fluency growth? 2) Do the effects of curricula on oral reading fluency outcomes differ by grade level? And, 3) does growth in achievement for lower-SES students vary depending on curriculum and grade level? Since curricula aim to follow state standards and federal guidelines for reading instruction, each should produce adequate reading skill growth for most students. However, it is possible that student growth may vary depending on grade and student characteristics (lower- or higher-SES) as a function of differing reading curricula.

Method

Participants: Reading First, Students, and Schools

The Reading First initiative. Reading First provides extensive funding for the improvement of schools demonstrating chronically low reading achievement, affecting approximately 264,254 students and 16,114 teachers within the state of Florida (Miller, 2007). Extensive funding is also provided for the improvement of schools demonstrating chronically low reading achievement, wherein monies are designated for teacher training, the purchase of scientifically based reading curricula, and assessment instruments (both norm-referenced and curriculum-based measures). Teacher training is intended to support the improvement of reading instruction, alongside a mandated 90-minute, language-arts instructional period, state-selected core curriculum adoption, and the use of assessments to guide instruction.

Students and schools. Each year, 10% of Florida Reading First schools (38 in this sample) are randomly selected to participate in site visits aimed at monitoring implementation of

Reading First. Students included in this study were a randomly selected sample of 10% of those who attended Reading First schools during the 2005–2006 school year; specifically, 9,993 first-grade students in 942 classrooms, 9,869 second-grade students in 962 classrooms, and 10,141 third-grade students in 954 classrooms. Students were not significantly different from sample statistics of Reading First schools overall. Student demographic data is presented in Table 1.

Measures

Oral reading fluency. Measures of oral reading fluency (ORF) are highly related to overall reading (Tindal & Marston, 1996), and correlate positively with standardized measures of reading achievement (Crawford, Tindal, & Stiver, 2001; Good, Simmons, & Kame'enui, 2001). ORF is often touted as the best overall indicator of reading proficiency for students in the early stages of learning to read (Shinn, Good, Knutson, & Tilly, 1992). Moreover, this measure is used throughout Florida, even in schools that do not receive Reading First funds, and is highly predictive of students' performance on state-mandated high-stakes assessments (Buck & Torgesen, 2006; Schatschneider et al., 2004). Test–retest reliability for the ORF assessment ranged from 0.92 to 0.97 across grade levels (Shaw & Shaw, 2002). For these reasons, ORF was chosen to represent reading achievement in this study.

Children's ORF was measured as the number of CRW/min. During the assessment, children were asked to read three passages for 1 minute each. If a student correctly said the word, the word was counted as correct. Misread words were considered to be incorrect and were not counted in the total words read correctly in 1 minute. Rates for individual passages were computed by calculating the number of CRW/min. The median score of three randomly selected passages was used to obtain an ORF score for the assessment. See Table 2 for fall and spring ORF score averages by curriculum and grade level. Passages were drawn from the Dynamic Indicators of Basic Early Literacy Skills (DIBELS) progress monitoring materials at first-, second-, and third-grade levels (Good & Kaminski, 2002). Assessments were conducted by reading coaches who were trained by Florida Reading First assessment teams and staff. Students were assessed at four time points during the school year (September, December, February, and April).

Florida Reading First benchmarks for ORF measures were designed to identify students at risk for reading failure by the end of the year and, for the purposes of this study, are used as benchmarks for adequate achievement in ORF. Risk levels were established by the authors of DIBELS reading fluency tests (Hasbrouck & Tindal, 1992; Fuchs, Fuchs, Hamlett, Walz, & Germann, 1993) and were revised by Florida Reading First to include an above-average reading benchmark. Benchmark levels include high, medium, and low risk for reading failure as well as above-average ORF achievement at each grade level. The above-average benchmark level was added to DIBELS norms for Florida Reading First schools as a normative indicator of high ORF skills at each assessment time point. (For a complete table see http://www.fcrr.org/pmrn/docs/dibles_risklevels_k3_0607_0809.pdf.) Indicators of risk are referred to in this study as achievement benchmarks to indicate if students are achieving at typical grade levels (i.e., low risk for reading difficulties). Adequate achievement at final assessment (spring) is designated as 40 CRW/min for first grade, 90 CRW/min for second grade, and 110 CRW/min for third grade.

Curricula

Schools in this study used one of six core reading curricula, including Open Court, Reading Mastery, Harcourt, Houghton Mifflin, Scott Foresman, and Success for All. (See Table 1 for distributions among schools.) For adoption eligibility in Florida Reading First schools, curricula

Table 1

Descriptive Statistics for Student Demographics and Achievement in Grades 1–3

Grade	Curriculum	% Gender		% Free or reduced lunch status			
		n	Male	DNQ	Free	RED	DNA
First		9993					
	Open Court	332	55	8	76	12	4
	Reading Mastery	727	46	3	63	10	24
	Harcourt	4499	52	7	60	11	22
	Houghton Mifflin	2067	53	4	77	10	9
	Scott Foresman	2078	51	2	65	10	23
	Success for All	290	54	6	62	10	22
Second		9869					
	Open Court	330	53	5	76	11	8
	Reading Mastery	705	52	1	62	10	27
	Harcourt	4346	53	7	58	13	22
	Houghton Mifflin	2121	53	4	74	11	11
	Scott Foresman	2082	52	4	63	9	24
	Success for All	285	55	7	62	12	19
Third		10141					
	Open Court	275	54	8	72	11	9
	Reading Mastery	656	52	26	60	11	3
	Harcourt	4551	52	23	56	13	8
	Houghton Mifflin	2307	52	10	74	12	4
	Scott Foresman	2074	52	24	62	10	4
	Success for All	278	52	20	60	13	7

Note. DNQ denotes students who applied, but did not quality. RED denotes students who applied and qualified for reduced-price lunches. DNA denotes students who did not apply for free or reduced-price lunch

Table 2

Descriptive Statistics for Oral Reading Fluency by Grade

	September ORF		April ORF	
	M	*SD*	*M*	*SD*
First grade				
All students	18	20	50	32
Higher-SES	25	16	63	35
Lower-SES	16	17	46	30
Open Court	14	17	45	30
Reading Mastery	22	20	57	32
Harcourt	18	20	51	32
Houghton Mifflin	16	19	44	31
Scott Foresman	19	21	52	32
Success for All	19	23	50	34
Second grade				
All students	53	31	89	35
Higher-SES	63	34	77	34
Lower-SES	50	29	62	30
Open Court	51	26	89	35
Reading Mastery	55	30	88	35
Harcourt	54	31	88	36
Houghton Mifflin	51	31	90	35
Scott Foresman	54	31	89	36
Success for All	56	30	92	33
Third grade				
All students	73	34	101	35
Higher-SES	85	34	113	35
Lower-SES	69	32	97	34
Open Court	78	32	100	35
Reading Mastery	79	32	109	34
Harcourt	74	34	103	35
Houghton Mifflin	68	34	96	36
Scott Foresman	73	33	102	34
Success for All	79	32	101	35

were required to meet standards of scientifically based reading research. Unfortunately, complete and thorough curriculum reviews, including intensity and frequency of use of curriculum components, are not available for Florida curricula. However, limited reviews conducted by the Florida Center for Reading Research (FCRR) during the 2002–03 Florida reading curriculum adoption indicated that all core curricula in our analyses met standards of scientifically based research, although detailed evaluations of curriculum components and frameworks were not available (FCRR, 2006). Notably, Success for All was reviewed by FCRR as an intervention program and was found to align with the five components of reading, which are central to the Reading First initiative (FCRR, 2002). Curricula were generally recognized as meeting the requirements of scientifically based reading research based on their inclusion of phonemic awareness, phonics, fluency, comprehension, and vocabulary instruction.

The following curriculum summaries, presented below in alphabetical order, are based on a limited review conducted for the purposes of describing curricula in this study. These helped to uncover the general framework of instruction for each program, and the aims of the resources, based on the sequence of instruction during the first unit of the year. Because this study occurred during the last year of the reading adoption, complete component reviews were not undertaken. New adoptions were slated to be made in the following year and reviewing them was beyond the scope of this study. However, depending on study results, such reviews may be warranted for future application in research and instructional decision making.

All curricula included levelized or decodable readers, which provided teachers with reading materials at multiple levels within a given grade level of the core curriculum. In this way, students could be provided text at their skill level. Programs also included accommodations for struggling students as well as those with special learning needs. While some curricula had similar frameworks and others differed in their approach, overall, phonics, phonemic awareness, fluency, comprehension, and vocabulary are described as part of the weekly plans for instruction provided for teachers in the teachers' manual.

Harcourt. Organized by weekly objectives within theme units, Harcourt lessons are divided into the following segments: Oral Language, Skills and Strategies, Reading, and Language Arts. For first and second grade, each daily lesson plan begins with Oral Language, typically comprised of a shared literature reading opportunity followed by phonemic awareness activities which focus on a particular skill each week. Lessons usually include the introduction of a new skill (letter or sound), followed by a spelling activity and high-frequency word practice. Next, Skills and Strategies are introduced, where comprehension skills are taught through the use of a basal and/or decodable book. Fluency and independent reading are also taught at this time. Language Arts activities of writing and grammar aim to introduce writing strategies and allow students a chance to practice each strategy within a given prompt. Grammar instruction, also skill-based, is followed by a daily language activity. Third grade follows a similar format for instruction, while an emphasis on phonics and phonemic awareness, although still present, is moderated by greater attention to vocabulary.

Houghton Mifflin. Houghton Mifflin lesson plans include the following aims: Learning to Read, Reading, Word Work, and Writing and Oral Language (http://www.educationplace.biz/rdg/hmr06/). These components serve as a framework for daily and weekly activities. For first grade, Learning to Read offers phonemic awareness, phonics, and comprehension instruction and practice. This involves the introduction of big books and stories from anthologies, which focus on a particular phonics skill and high-frequency words. Next, com-

prehension and phonics skills are practiced, followed by the reading of a decodable book exemplifying skills previously taught. Word Work includes spelling and phonics, high-frequency words, and vocabulary. Writing and Oral Language embodies opportunities such as shared, interactive, or independent writing, and on some days grammar and listening comprehension are introduced.

For second grade, instruction types (Reading, Word Work, or Writing and Oral Language) remain the same; however, Learning to Read is replaced with Reading, which removes the focus from phonemic awareness, placing it instead on opportunities for instruction and practice of comprehension skills. During Writing and Oral Language, daily prompt and grammar lessons are used in addition to daily language practice. For third grade, phonics instruction is replaced by comprehension strategy instruction; fluency practice and independent reading also make up this time.

Open Court. Organized by theme units with scripted daily lessons, Open Court daily plans are segmented by topic, including Preparing to Read, Reading and Responding, Inquiry, and Language Arts (https://www.sraonline.com/). In first grade, Preparing to Read involves phonemic awareness, phonics, and dictation of words and sentences. Reading and Responding instructs teachers to read aloud and ask questions about the story, and use comprehension and vocabulary strategies to support the text. Inquiry in first grade extends student knowledge about story content and helps students make connections to math and art. Language Arts involves word analysis, vocabulary, writing strategies, and grammar lessons.

Second-grade activities during Preparing to Read include text reading, phonic activities, and practice using words and sentences which exemplify skills taught. Reading and Responding asks students to activate prior knowledge about what they will read together, and focus on its vocabulary. Inquiry allows students to ask questions about the content presented in daily readings, including strategies to organize facts and information. Language Arts instruction focuses on more complex phonics skills, writing strategies, and daily grammar lessons.

In third grade, the same general format of second grade is followed, with more discussion of stories and comprehension strategies, followed by word study rather than phonics. Key differences from second grade include a greater use of critical thinking and comprehension strategies during Reading and Responding. Inquiry activities require students to think and apply information learned from text readings, while Language Arts instruction maintains a similar focus, providing daily writing and grammar lessons.

Reading Mastery. Reading Mastery lessons are scripted and characterized by a direct style of instruction (https://www.sraonline.com/). For first grade, daily instruction begins with a review of previously introduced phonics and phonemic awareness skills. Next, an opportunity to use those sounds within the context of real words is presented as a precursor to later inclusion in story readings. Sight words, which later appear in texts for practice reading, are introduced and identified as irregular or regular. Once the items specific to daily stories are introduced, the story is read. Students participate in multiple readings of text for each lesson with specific aims, such as reading to practice decoding skills and reading for comprehension. Last, activities for further review of daily concepts are offered. Expansion activities for reading are also introduced after each lesson, as an option for the teacher, and are often integrated with subjects aside from reading or language arts. Other daily lesson plan components include Spelling and Language.

Second and third grade follow a similar framework, with different emphases. Like first graders, second and third graders are presented with phonics, word families, or word

segments; however, instead of a focus on high-frequency words, vocabulary related to the daily story is introduced and reviewed. Irregular and regular words are also taught and reviewed. Daily practice reads are conducted in the same format as first grade, followed by comprehension activities and strategies. Practice activities also are offered, along with seatwork for additional independent learning. Spelling and language lessons for each day are followed by literature lessons and independent readers, as an option for further extension.

Scott Foresman. The following components make up Scott Foresman lessons: Reading, Oral Language, Writing, and Self Selected Reading/Read Alouds (http://www.sfreading.com/). In first grade, daily texts for Reading activities vary day by day so that students read both basal stories and independent readers. Vocabulary follows a weekly theme, related to the overall nature of the readings. Speaking, Listening and Viewing activities are offered next, allowing time for students to engage in listening comprehension and oral expression. Writing lessons involve different modes of writing (e.g., shared, independent) and are coupled with grammar and language lessons. Daily Self Selected Reading and Read Alouds are included, along with plans of how to engage students, what types of books to use, and how to discuss readings with students.

Success for All. Success for All is an entire school reform curriculum, which requires participation from administration and parents as well as teachers and students (FCRR, 2002; http://successforall.com/). The materials are not organized by grade, but rather by assessed ability level. They consist of three parts: KinderRoots, ReadingRoots, and Reading Wings. KinderRoots and ReadingRoots begin with a read-aloud time. This is followed with an opportunity for shared reading of decodable books, which is then succeeded by explicit instruction in phonics, phonemic awareness, and vocabulary instruction, depending on the nature of the lesson and the grade level. Opportunities for language development then follow, including writing instruction.

Fidelity of Curriculum Use

To ensure compliance with Reading First guidelines, randomly selected schools and classrooms were observed by Reading First observation teams during either the fall or spring of the 2005–06 school year. A subset of schools included in this study was randomly chosen to participate in site visits and serve as an indicator of fidelity. These visits included interviews with principals, reading coaches, and teacher focus groups who were asked to report the level to which they felt the core was a part of their reading instruction. Specifically, reading coaches were asked to report the role of the core curriculum at their school, to which 36 out of 38 schools reported that it was a significant or principal part of instruction. Even though the use of the curricula was judged important in the two remaining schools, for some grade levels, more components of the core were used than in others. Reading coaches reported that this was due to strengths and weaknesses of different curricula. School personnel who participated in random site visits were also asked whether other supplemental materials were used to accomplish instructional goals aside from the core curriculum, and they noted that other resources were used in addition to the core. All curricula were supplemented for students who were struggling under the regular curricula. These intervention programs varied across schools, but within schools and curricula, students were taught using similar programs. These include Great Leaps, Waterford/Successmaker, LeapFrog Accessories, Voyager, Early Success, Read, Write, and Type, and Earobics. Based on interviews and observations, one particular intervention program did not seem to be more frequently used than any other.

In some cases, schools using a curriculum other than Reading Mastery used the programs

to supplement their instruction for struggling readers. In some cases, school curriculum differed by grade. However, since the analyses were conducted based on the curriculum assignment for each student, it is not possible to examine this variable as a mediator in this study. Supplementary materials are reported at the school level rather than the student level and therefore cannot be used in analytic models. Additionally, a stated responsibility of the reading coach at each Reading First school was to ensure fidelity and implementation of the core curriculum, as funds for Reading First schools were dependent on compliance with terms of the program. All schools were expected to maintain fidelity to their chosen core and, based upon observations, most classrooms were found to be utilizing the core on the particular day of observation.

Socioeconomic Status

Students' eligibility for free or reduced-price lunch (lower-SES) was used as a proxy for students' SES status. Lower-SES students were designated as those who applied and were eligible for free or reduced lunches. Students who did not apply or applied but were not eligible were not considered to be lower-SES in our analyses. To qualify for free or reduced-price lunches students had to live in families with lower to very low incomes, falling below 135% of the poverty level for free lunch ($26,112 for a family of 4) and 185% for reduced-price lunch ($35,797 for a family of 4; http://aspe.hhs.gov/poverty/05poverty.shtml). Percentages of students identified as lower-SES were consistent across grade levels and among curricula (see Table 1).

Model Testing and Analyses

To accommodate the nested nature of the data (repeated assessments of ORF nested within students and students nested in classrooms using a particular curriculum), ORF growth curves were modeled using Hierarchical Linear Modeling (HLM; Raudenbush & Bryk, 2002). Using HLM, estimating mean growth trajectories allowed the examination of how children's SES status affected growth over the school year (7 months). All analyses were centered at the spring assessment time point (April), thus intercepts represent fitted mean end-of-year ORF scores in terms of CRW/min.

Models for research questions I and II: What are the effects of different core curricula on children's reading fluency growth? Do the effects of curricula on ORF outcomes differ by grade level?

First, an unconditional model was tested at each grade level in order to examine whether ORF achievement varied randomly at the student and classroom level. Significant variance components from these analyses ($p<0.01$) were used to confirm that multilevel analyses were necessary to take into account random variance across students and classrooms, since ORF varied randomly across both. At each grade level, model comparison tests were performed and models with both linear (Time) and quadratic ($Time^2$) terms provided a significantly better fit than the linear term only (p value <0.01 for each grade level). Classroom variance (pseudo R^2) increased from the linear to the quadratic model for first- (0.67, 0.76) and third-grade models (0.65, 0.89), while second-grade models with only the linear term (0.97) were a better fit than those which included a quadratic term (0.67). Since model comparison tests indicated that both quadratic models were most explanatory, and that variance explained was greater for the quadratic model at first and third grade, these were applied to each grade level.

Level-1:

$$Y_{tij}=\pi_{0ij} + \pi_{1ij}(Time_{tij}) + \pi_{2ij}(Time^2_{tij}) + e_{tij}$$

Level-2:

$$\pi_{0ij}=\beta_{00j}+r_{0ij}$$
$$\pi_{1ij}=\beta_{10j}+r_{1ij}$$
$$\pi_{2ij}=\beta_{20j}$$

Level-3:
$B_{00j}=\gamma_{001}$ (Open Court) $+\gamma_{002}$ (Reading Mastery) $+\gamma_{003}$ (Harcourt) $+\gamma_{004}$ (Houghton Mifflin) $+\gamma_{005}$ (Scott Foresman) $+\gamma_{006}$ (Success for All) $+u_{00j}$

$\beta_{10j}=\gamma_{101}$ (Open Court) $+\gamma_{102}$ (Reading Mastery) $+\gamma_{103}$ (Harcourt) $+\gamma_{104}$ (Houghton Mifflin) $+\gamma_{105}$ (Scott Foresman) $+\gamma_{106}$ (Success for All) $+u_{10j}$
$\beta_{20j}=\gamma_{201}$ (Open Court) $+\gamma_{202}$ (Reading Mastery) $+\gamma_{203}$ (Harcourt) $+\gamma_{204}$ (Houghton Mifflin) $+\gamma_{205}$ (Scott Foresman) $+\gamma_{206}$ (Success for All)

Where Y_{tij} is the fitted spring ORF score at time t for child i in classroom j and is a function of the linear ($Time_{tij}$; $\gamma_{101} \ldots \gamma_{106}$) and quadratic ($Time^2_{itj}$; $\gamma_{201} \ldots \gamma_{206}$) growth terms for each curriculum added to the intercept or fitted mean for a particular curriculum ($\gamma_{001} \ldots \gamma_{006}$). Residuals were assumed to be normally distributed.

The degree to which intercept, slope, and quadratic terms varied across students and classrooms was tested in order to examine whether effects should be random or fixed. For all grade levels, intercept and slope varied randomly across students and classrooms (curricula) ($p<0.01$), while the quadratic term did not ($p=0.50$). Therefore, intercept and slope were allowed to vary randomly and quadratic was assigned to be a fixed effect.

After the suitability of growth terms was examined, growth across curricula was examined using four ORF scores modeled at Level 1 (September, December, February, April) and curriculum at Level 3 (Table 3; Figure 1, top; Figure 2, top; Figure 3, top). Curricula were dummy coded (1 or 0) for individual modeling of growth. Parameters of intercept, slope, and quadratic were used to describe growth as well as acceleration or deceleration. At Level 3, the intercept was removed to examine individual growth curves for each curriculum. This strategy removes the need for a reference group and instead examines whether model coefficients for each curriculum (intercept, slope, and quadratic) are significantly different from zero rather than a particular comparison group. Pairwise comparisons were performed in order to examine how each intercept, slope, and quadratic coefficient differed across curricula. The linear step-up method was used to correct for multiple comparisons and minimize the false discovery rate of significant effects (Benjamini and Hochberg, 1995). Effect sizes for growth models with the inclusion of SES were calculated by subtracting the lower of two scores from the highest, then dividing that difference by the baseline standard deviation of ORF scores for each grade respectively. Cohen's cutoffs (small=0.2; medium=0.5; large=0.8) were applied when discussing the nature of each effect size (Cohen, 1988).

Models for research question III: Does growth in achievement for lower-SES students vary depending on curriculum and grade level?

Upon inspection of growth curves and effect sizes for growth models at each grade level, it seemed that more often than not, curricula with the lowest (fitted) scores in April were also those which had the lowest (fitted) initial ORF scores at the beginning of the year. Therefore, models were revised in order to control for students' initial status while also examining the influence of SES within the same population. Growth trajectories over the school year were described by the remaining three scores (late fall, winter, and spring) at Level 1, fall ORF scores (grand-mean centered) and SES (group-mean centered) were entered at the child level (Level 2), and curricula were entered at Level 3. Due to a lack of degrees of freedom to model the quadratic trend, only slopes were modeled (Raudenbush & Bryk, 2002). Harcourt served as the fixed reference group, since the majority of students in the sample were using this curriculum. Thus, the coefficients for each curriculum represent their fitted mean difference from the Harcourt students' performance.

Table 3

Grade 1–3 HLM Overall Growth Model for Spring Oral Reading Fluency

ME	First grade			Second grade			Third grade		
	CO	*SE*	*p* value (*df*)	CO	*SE*	*p* value (*df*)	CO	*SE*	*p* value (*df*)
OC	44.16	2.43	<0.01(936)	80.63	3.96	<0.01(956)	92.17	4.42	<0.01(948)
RM	54.99	2.73	<0.01(936)	90.74	3.51	<0.01(956)	104.98	3.44	<0.01(948)
HC	50.06	0.85	<0.01(936)	85.66	1.06	<0.01(956)	98.76	1.10	<0.01(948)
HO	42.52	1.26	<0.01(936)	79.34	1.81	<0.01(956)	92.30	1.58	<0.01(948)
SF	51.44	1.15	<0.01(936)	84.00	1.69	<0.01(956)	97.69	1.66	<0.01(948)
SA	49.92	4.00	<0.01(936)	90.84	3.82	<0.01(956)	96.69	4.33	<0.01(948)
Linear									
OC	6.97	0.51	<0.01(936)	6.95	0.45	<0.01(956)	0.76	0.54	0.17(948)
RM	5.36	0.29	<0.01(936)	6.46	0.51	<0.01(956)	2.04	0.45	<0.01(948)
HC	6.03	0.12	<0.01(936)	6.11	0.13	<0.01(956)	3.07	0.15	<0.01(948)
HO	5.17	0.21	<0.01(936)	5.99	0.17	<0.01(956)	2.98	0.23	<0.01(948)
SF	6.14	0.18	<0.01(936)	5.63	0.21	<0.01(956)	2.90	0.23	<0.01(948)
SA	5.18	0.38	<0.01(936)	5.67	0.44	<0.01(956)	2.17	0.44	<0.01(948)
Quadratic									
OC	0.34	0.06	<0.01(39789)	0.30	0.05	<0.01(39308)	−0.33	0.06	<0.01(40391)
RM	0.05	0.04	0.19(39789)	0.10	0.05	0.07(39308)	−0.32	0.06	<0.01(40391)
HC	0.17	0.01	<0.01(39789)	0.15	0.02	<0.01(39308)	−0.15	0.02	<0.01(40391)
HO	0.17	0.02	<0.01(39789)	0.19	0.02	<0.01(39308)	−0.14	0.03	<0.01(40391)
SF	0.20	0.02	<0.01(39789)	0.12	0.02	<0.01(39308)	−0.19	0.03	<0.01(40391)
SA	0.11	0.05	0.04(39789)	0.08	0.06	0.18(39308)	−0.16	0.06	0.01(119374)

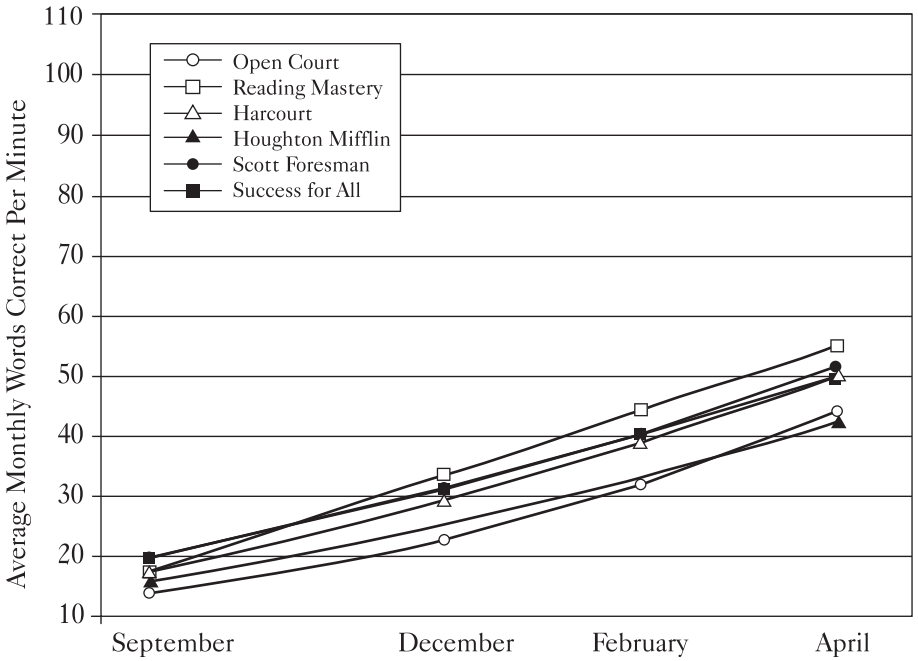
Random effects

	First grade			Second grade			Third grade		
	Var	X^2	*p* value	Var	X^2	*p* value	Var	X^2	*p* value
Class	176.2	323.4	<0.01	363.1	4084.4	<0.01	777.3	528642.7	<0.01
Linear	52.4			71.7			71.9		
Child	814	225845.9	<0.01	916.5	18239.4	<0.01	858.2	22541.5	<0.01

Note. ME = main effect, CO = coefficient, *SE* = standard error, and *df* = degrees of freedom. OC = Open Court, RM = Reading Mastery, HC = Harcourt, HO = Houghton Mifflin, SF = Scott Foresman, and SE = Success for All.

Figure 1

First-Grade Overall ORF Growth



First-Grade SES Differences at Spring

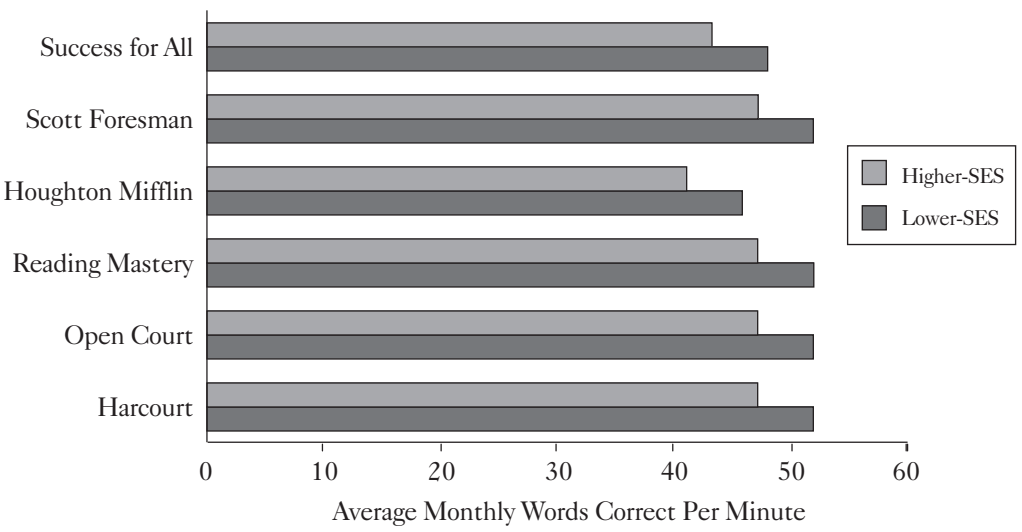
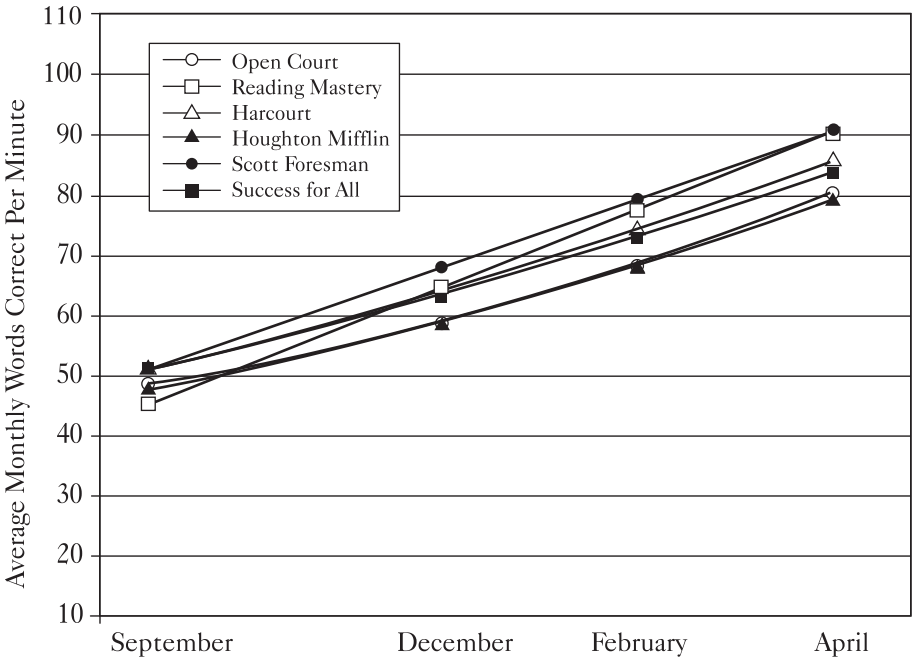


Figure 2

Second-Grade Overall ORF Growth



Second-Grade SES Differences at Spring

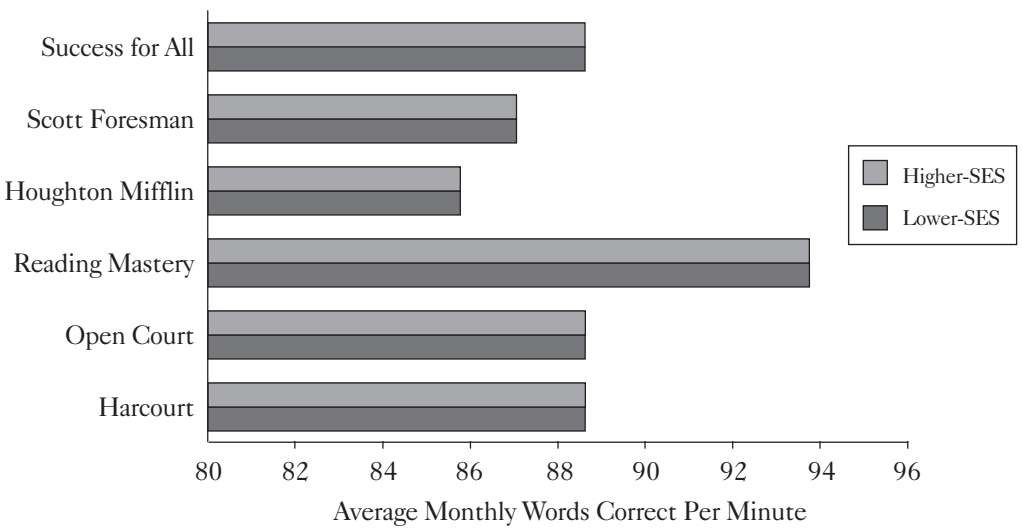
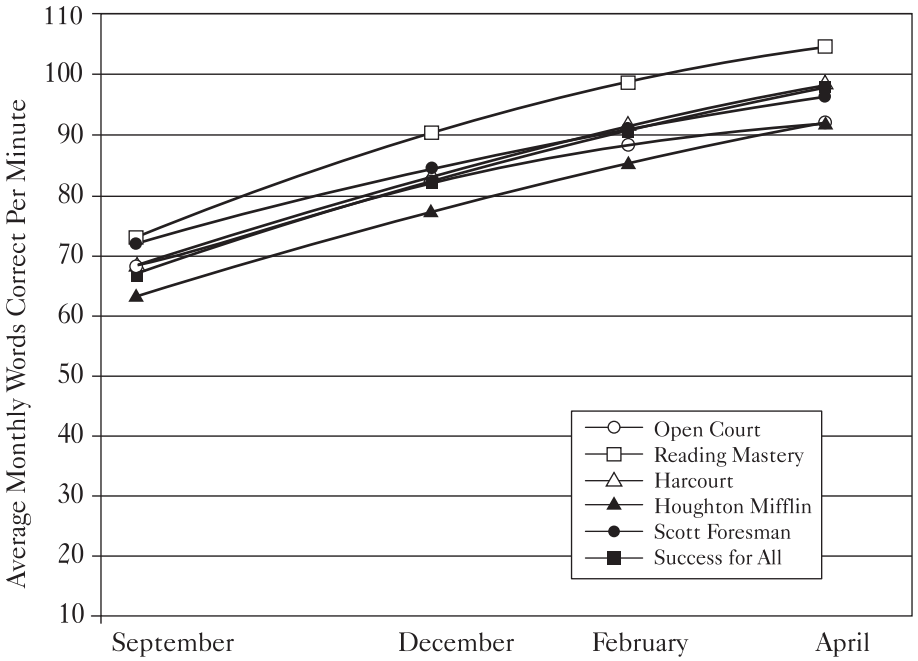
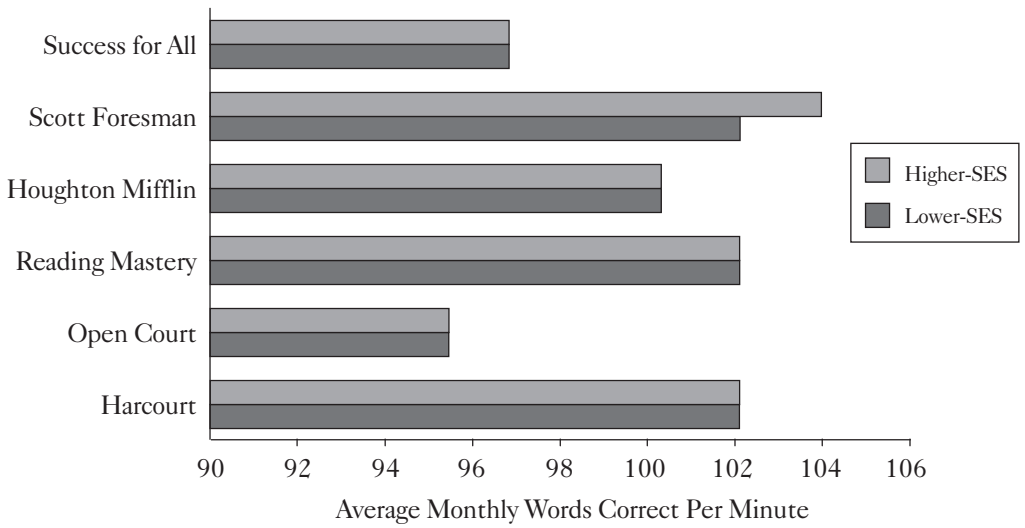


Figure 3

Third-Grade Overall ORF Growth



Third-Grade SES Differences at Spring



Level-1:

$Y_{tij}=\pi_{0ij}+\pi_{1ij}(\text{Time}_{tij})+e_{tij}$

Level-2:

$\pi_{0ij}=\beta_{00j}+\beta_{01j}(\text{Initial ORF})+\beta_{02}(\text{FRL})+r_{ij}$

$\pi_{1ij}=\beta_{10j}+\beta_{11j}(\text{Initial ORF})+\beta_{12}(\text{FRL})$

Level-3:

$\beta_{00j}=\gamma_{000}+\gamma_{001}(\text{Open Court})+\gamma_{002}(\text{Reading Mastery})+\gamma_{003}(\text{Houghton Mifflin})+\gamma_{004}(\text{Scott Foresman})+\gamma_{005}(\text{Success for All})+u_{00j}$

$\beta_{01j}=\gamma_{010}$

$\beta_{02j}=\gamma_{020}+\gamma_{021}(\text{Open Court})+\gamma_{022}(\text{Reading Mastery})+\gamma_{023}(\text{Houghton Mifflin})+\gamma_{024}(\text{Scott Foresman})+\gamma_{025}(\text{Success for All})$

$\beta_{10j}=\gamma_{100}+\gamma_{101}(\text{Open Court})+\gamma_{102}(\text{Reading Mastery})+\gamma_{103}(\text{Houghton Mifflin})+\gamma_{104}(\text{Scott Foresman})+\gamma_{105}(\text{Success for All})$

$\beta_{11j}=\gamma_{110}$

$\beta_{12j}=\gamma_{120}+\gamma_{121}(\text{Open Court})+\gamma_{122}(\text{Reading Mastery})+\gamma_{123}(\text{Houghton Mifflin})+\gamma_{124}(\text{Scott Foresman})+\gamma_{125}(\text{Success for All})$

Where Y_{tij} is the fitted spring ORF score at time *t* for child *i* in classroom *j* and is a function of the fitted mean for the sample (γ_{000}) plus the fitted mean linear growth (γ_{100}) plus the coefficient or effect of each curriculum ($\gamma_{000} \ldots \gamma_{005}$) plus linear growth coefficients for each curriculum ($\gamma_{101} \ldots \gamma_{105}$), controlling for initial ORF score (γ_{010} & γ_{110}), as well as the interaction between SES (FRL) and each curriculum affecting outcome ($\gamma_{020} \ldots \gamma_{025}$) and growth ($\gamma_{120} \ldots \gamma_{125}$).

Results

Intraclass Correlations for the Growth Model

The unconditional first-grade model revealed an intraclass correlation of 0.19, which is the proportion of total variance between classrooms. Intraclass correlations are calculated by dividing the sum of the variance by the classroom-level variance. Variance components suggest that 77% of the variance fell between children, taking into account classroom variance. Variance between students is calculated by dividing the child-level variance by the total variance of a model, which in this case includes ORF and TIME. The unconditional second-grade model revealed an intraclass correlation of 0.21.Variance components suggest that 67% of the variance fell between children, taking into account classroom variance. The unconditional third-grade model revealed an intraclass correlation of 0.23. Variance components suggest that 65% of the variance fell between children, taking into account classroom variance.

Growth Model with SES

The unconditional first-grade model (three measures of ORF) revealed an intraclass correlation of 0.15. Variance components suggest that 74% of the variance fell between children, taking into account classroom variance. The unconditional second-grade model (three measures of ORF) revealed an intraclass correlation of 0.24. Variance components suggest that 61% of the variance fell between children, taking into account classroom variance. The unconditional third-grade model (three measures of ORF) revealed an intraclass correlation of 0.26. Variance components suggest that 70% of the variance fell between children, taking into account classroom variance.

Research questions I and II: What are the effects of different core curricula on children's reading fluency growth? Do the effects of curricula on ORF outcomes differ by grade level?

On average, first-grade ORF scores increased from September to April, characterized by growth that was generally linear with some acceleration. Overall, students were at or above ORF benchmarks of adequate achievement at the April assessment (see Figure 1 top; 40 CRW/min). Reading Mastery students were, on average, reading 55 CRW/min at the final first-grade assessment time point (April) followed by Scott Foresman (51), Harcourt (50), Success for All (50), Open Court (44), and Houghton Mifflin (43). On average, Houghton Mifflin scores were significantly lower than those observed for Reading Mastery, Harcourt, and Scott Foresman students (p values <0.01) and Open Court students had significantly lower scores than Reading Mastery and Harcourt (p values: <0.01). In terms of growth, linear coefficients (change in WRC/min between assessment time points) did not significantly differ across curricula (range of 5.17–6.97), while quadratic trends (acceleration or deceleration of correctly read words per minute between assessment time points) varied (Open Court, 0.34; Scott Foresman, 0.20; Harcourt and Houghton Mifflin, 0.17; Success for All, 0.11; Reading Mastery, 0.05). For Open Court, acceleration was significantly greater than any other curricula (0.34; p values <0.02), and Reading Mastery was significantly lower (0.05; p values <0.01) when compared to all other curricula with the exception of Success for All.

For second grade, on average, ORF growth trajectories were similar in linearity and acceleration to those observed for first-grade students. Growth curve plots are provided in Figure 2. In terms of final status, by the last assessment time point (April), only students in Reading Mastery and Success for All classrooms met the adequate achievement benchmark (90 CRW/min). While this suggests substantial differences between curricula, few programs differed significantly at the final assessment time point. Reading Mastery and Success for All students had the highest fitted ORF scores at the final assessment time point (April; 91) while students in other curricula scored slightly lower (Harcourt, 86 CRW/min; Scott Foresman, 84 CRW/min; Open Court, 81 CRW/min; and Houghton Mifflin, 79 CRW/min). Only Houghton Mifflin students' (79) scores were significantly different than those observed for other curricula, as Houghton Mifflin students were significantly lower than students using Reading Mastery or Success for All (p values <0.01). Linear growth coefficients (monthly change in CRW/min) across curricula did not differ at the end of the year (range of 5.63–6.95 CRW/min); however, significantly different rates of acceleration for ORF were detected. Students using Open Court (0.05) exhibited growth that accelerated significantly more from September to April (p values <0.01) in comparison to all curricula except Houghton Mifflin.

While first- and second-grade growth curves were characterized by acceleration from September to April, third graders demonstrated trends of deceleration (Figure 3, top), which indicates that student ORF growth/month was slowing over the school year instead of increasing. On average, and regardless of curriculum, students did not meet the adequate achievement benchmarks at the end of the school year (110 CRW/min). At the April assessment, Reading Mastery students had the highest ORF scores (105 CRW/min) while other curricula ranged from 92 CRW/min (Open Court) to 99 CRW/min (Harcourt). Significant differences were detected between students using Houghton Mifflin and those using Harcourt and Reading Mastery. Although linear growth (change in CRW/min between assessment time points) for first and second graders did not vary across curricula, in third grade significant differences were apparent. Results suggest that Open Court students had significantly less linear growth (0.76 CRW/min) than Harcourt (3.07 CRW/min), Houghton Mifflin (2.98 CRW/min), and Scott Foresman (2.90 CRW/min) (p values <0.01). Acceleration parameters (quadratic trend) also varied significantly for Open Court (0.06) and

Reading Mastery (0.06) students who experienced more deceleration in CRW/min than Harcourt (0.02) and Houghton Mifflin (0.03) (p values<0.01).

Research question III: Does growth in achievement for lower-SES students vary depending on curriculum and grade level?

In first grade, higher-SES students using Open Court, Reading Mastery, and Scott Foresman did not significantly differ from higher-SES Harcourt students (52 CRW/min at April) while higher-SES Houghton Mifflin (46 CRW/min; effect size=0.20) and Success for All students (48 CRW/min; effect size=0.16) scored significantly lower. Lower-SES students using Open Court, Reading Mastery, Scott Foresman, and Harcourt, scored, on average, 47 CRW/min while lower-SES Houghton Mifflin (41 CRW/ min; effect size=0.20) and Success for All (43 CRW/min; effect size=0.16) students scored significantly lower. ORF growth for first grade, higher-SES students using Open Court, Reading Mastery, and Scott Foresman did not significantly differ from higher-SES Harcourt students (monthly increase of 5.79 CRW/min) while higher-SES students using Houghton Mifflin (4.89 CRW/min monthly increase) and Success for All (4.99 CRW/min monthly increase) experienced slightly less growth. Lower-SES students using Open Court, Reading Mastery, Scott Foresman, and Harcourt grew slightly less (5.05 CRW/min monthly increase) than their higher-SES counterparts. Coupled with lower ORF achievement, lower-SES students using Houghton Mifflin and Success for All experienced slighter score increases than students using Harcourt (4.89 and 4.99 CRW/min monthly increase). Despite differences dependent on curricula and SES status, on average, students reached first-grade ORF benchmarks upon final assessment (Figure 1, bottom; Table 3).

For second grade, higher-SES Open Court and Success for All students had April ORF scores that were not significantly different from higher-SES Harcourt students (89 CRW/min). In contrast, higher-SES Reading Mastery students scored significantly higher than Harcourt students (94 CRW/min; effect size=0.14) while higher-SES Houghton Mifflin (86 CRW/min; effect size=0.08) and Scott Foresman students (87 CRW/min; effect size=0.04) scored significantly lower than those using Harcourt. Differences between lower- and higher-SES students were detected among those using Open Court. Higher-SES students scored, on average, 89 CRW/min, whereas lower-SES students were scoring 85 CRW/min (effect size=0.11). Overall, for both lower- and higher-SES students, only those using Reading Mastery met ORF achievement benchmarks by the last assessment time point (Figure 2, bottom; Table 4).

ORF growth for higher-SES second graders did not significantly differ between Harcourt (5.69 CRW/min), Open Court, Reading Mastery, and Success for All. Significant differences were detected for higher-SES students using Houghton Mifflin (5 CRW/min) and Scott Foresman (5 CRW/min). Within curricula, no significant differences in growth between lower and higher-SES students were detected. Therefore, on average, second grade, lower-SES students were growing at the same rate as their higher-SES peers who used the same curricula. In third grade, higher-SES Reading Mastery and Scott Foresman students did not differ significantly from those using Harcourt (102 CRW/min) while Open Court (95 CRW/min; effect size=0.19), Success for All (97 CRW/min; effect size=0.14), and Houghton Mifflin students (100 CRW/min; effect size=0.05) scored significantly lower than those using Harcourt. Concomitantly, lower-SES students using Scott Foresman scored slightly higher (104 CRW/min) than their higher-SES peers (102 CRW/min) who were also using Scott Foresman (effect size=0.05). Overall, third-grade ORF scores at the last assessment time point (April) indicate that students were approaching ORF achievement benchmarks

Table 4

Grade 1 HLM Model for Spring Oral Reading Fluency, Controlling for Oral Reading Fluency Initial Status and SES

Curriculum	Coefficient	*SE*	*df*	*p* value	Calculated score
Higher-SES intercept					
Harcourt	51.91	0.50	935	0.00	51.91
Open Court	−0.73	1.75	935	0.68	51.91
Reading Mastery	0.60	1.31	935	0.64	51.91
Houghton Mifflin	−6.01	0.89	935	0.00	45.90
Scott Foresman	−1.35	0.88	935	0.13	51.91
Success for All	−3.85	1.81	935	0.03	48.06
Adjustment on higher-SES intercept for lower-SES					
Harcourt	−4.73	0.80	9980	0.00	−4.73
Open Court	−5.61	3.99	9980	0.16	−4.73
Reading Mastery	2.48	1.71	9980	0.15	−4.73
Houghton Mifflin	2.39	1.46	9980	0.10	−4.73
Scott Foresman	1.55	1.34	9980	0.25	−4.73
Success for All	4.42	2.84	9980	0.12	−4.73
Higher-SES slope					
Harcourt	5.79	0.09	935	0.00	5.79
Open Court	0.38	0.34	935	0.27	5.79
Reading Mastery	−0.31	0.21	935	0.15	5.79
Houghton Mifflin	−0.90	0.16	935	0.00	4.89
Scott Foresman	−0.13	0.16	935	0.44	5.79
Success for All	−0.80	0.28	935	0.01	4.99
Adjustment on higher-SES slope for lower-SES					
Harcourt	−0.74	0.15	9980	0.00	−0.74
Open Court	−1.29	0.75	9980	0.08	−0.74
Reading Mastery	0.21	0.33	9980	0.53	−0.74
Houghton Mifflin	0.56	0.28	9980	0.05	−0.18
Scott Foresman	0.21	0.25	9980	0.41	−0.74
Success for All	0.42	0.48	9980	0.38	−0.74
Lower-SES on intercept	1.22	0.01	9980	0.00	
Lower-SES on slope	0.03	0.00	9980	0.00	

Random Effects	Variance	Chi square	*p* value
Class	49.80	2350.00	<0.01
Linear	53.52		
Child	313.72	71680.59	<0.01

Note. Calculated score denotes the adjusted mean or slope for each curriculum, depending on its significance from the reference group (Harcourt).

(110 CRW/min), but on average, and regardless of curricula, students did not reach this goal (Figure 3, bottom; Table 5).

In third grade, students using Harcourt grew, on average, 3.85 CRW/min/month. Only Open Court and Success for All differed significantly from this rate (2.2 and 2.82, respectively). Across curricula, no differences in growth were detected between lower- and higher-SES students (see Table 6).

Discussion

The purpose of this study was to investigate the effect of six core reading curricula on students' ORF growth while considering possible influences of students' grade and SES. Results for each grade were compared to achievement benchmarks for the end of the school year, which were set by the test authors and the State of Florida. In first grade, students were achieving adequate reading fluency skill growth and achieved adequate achievement (final benchmark=40 CRW/min) by the end of first grade; however, significant differences among curricula were detected. Those differences are discussed below. By the end of second grade, on average, students' reading skills fell only slightly below the benchmark of adequate achievement (90 CRW/min); although, in second grade, this depended on the core curriculum used in the classroom. By the end of third grade, on average, students did not meet set benchmarks (110 CRW/min) regardless of the curriculum their teachers used. It was hypothesized that ORF growth would be similar across curricula used in Reading First schools in the state of Florida. However, results suggest that for first-, second-, and third-grade students, ORF growth differed by curriculum as well as by grade level. Even controlling for students' initial ORF score and SES status, some differences among curricula and across grades remained.

Overall, students in the Reading Mastery curriculum demonstrated generally greater overall ORF growth than students in other curricula. Also, they more frequently met or exceeded benchmarks for adequate achievement in first, second, and third grade. In first grade, regardless of SES status, students generally met adequate achievement benchmarks. Among second graders, on average, only students using Reading Mastery and Success for All met benchmarks, while the lowest scores for students were among those using Houghton Mifflin. In third grade, on average, students did not reach the adequate achievement benchmark. However, Reading Mastery students came closest to the benchmarks because scores among these students were the highest across curricula.

Taken together, overall results indicate that some reading curricula seem to be associated with higher ORF scores. Specifically, students in Reading First schools and classrooms using Reading Mastery demonstrated stronger ORF skills when compared to students in classrooms using other curricula. This was particularly apparent in first grade, where effect sizes were greatest (0.44) for using Reading Mastery compared to students using other reading curricula. Reading First schools, by definition, serve a high proportion of children from lower-SES homes with weaker overall academic achievement. Extant research points to the need for greater amounts of explicit, basic reading skill instruction for students who start school with weaker skills (Foorman et al., 2006; Snow, Burns, & Griffin, 1998; Torgesen et al., 2001). It is possible that the highly scripted structure of Reading Mastery curriculum may provide relatively more explicit instruction than the other curricular series examined in this study. It is also possible that Reading Mastery was supplemented with other materials, which may have contributed to this finding. This factor could be a contributor to the results for any of the curricula investigated, as supported by evidence gathered at the participating schools. Additionally, the measured reading outcome

Table 5

Grade 2 HLM Model for Spring Oral Reading Fluency, Controlling for Initial Status and SES

Curriculum	Coefficient	*SE*	*df*	*p* value	Calculated score
Higher-SES intercept					
Harcourt	88.63	0.41	956	0.00	88.63
Open Court	−1.50	1.73	956	0.39	88.63
Reading Mastery	5.14	1.47	956	0.00	93.77
Houghton Mifflin	−2.84	0.76	956	0.00	85.78
Scott Foresman	−1.57	0.80	956	0.05	87.06
Success for All	−0.01	1.29	956	1.00	88.63
Adjustment from higher-SES intercept for lower-SES					
Harcourt	−1.08	0.58	9866	0.06	0.00
Open Court	−3.96	2.06	9866	0.05	0.00
Reading Mastery	−2.16	1.62	9866	0.18	0.00
Houghton Mifflin	0.57	1.16	9866	0.62	0.00
Scott Foresman	1.27	0.96	9866	0.19	0.00
Success for All	2.37	2.33	9866	0.31	0.00
Higher-SES slope					
Harcourt	5.69	0.09	956	0.00	5.69
Open Court	0.08	0.31	956	0.80	5.69
Reading Mastery	0.58	0.31	956	0.06	5.69
Houghton Mifflin	−0.35	0.15	956	0.02	5.33
Scott Foresman	−0.34	0.15	956	0.03	5.35
Success for All	−0.28	0.25	956	0.27	5.69
Adjustment from higher-SES slope for lower-SES					
Harcourt	−0.06	0.13	9866	0.62	0.00
Open Court	−0.69	0.45	9866	0.12	0.00
Reading Mastery	−0.25	0.34	9866	0.45	0.00
Houghton Mifflin	0.27	0.28	9866	0.35	0.00
Scott Foresman	0.13	0.20	9866	0.50	0.00
Success for All	0.76	0.54	9866	0.16	0.00
Lower-SES on intercept	1.00	0.01	9866	0.00	
Lower-SES on slope	0.01	0.00	9866	0.00	

Random Effects	Variance	Chi square	*p* value
Class	42.32	2587.97	<0.01
Linear	73.45		
Child	188.22	35611.75	<0.01

Note. Calculated score denotes the adjusted mean or slope for each curriculum, depending on its significance from the reference group (Harcourt).

Table 6

Grade 3 HLM Model for Spring Oral Reading Fluency, Controlling for Initial Status and SES

Curriculum	Coefficient	*SE*	*df*	*p* value	Calculated score
Higher-SES intercept					
Harcourt	102.16	0.38	948	0.00	102.16
Open Court	−6.70	1.68	948	0.00	95.46
Reading Mastery	1.40	1.26	948	0.27	102.16
Houghton Mifflin	−1.83	0.71	948	0.01	100.33
Scott Foresman	0.37	0.71	948	0.61	102.16
Success for All	−5.31	0.81	948	0.00	96.85
Adjustment from high SES intercept for lower-SES					
Harcourt	0.03	0.49	10137	0.96	0.00
Open Court	0.64	2.81	10137	0.82	0.00
Reading Mastery	1.06	1.50	10137	0.48	0.00
Houghton Mifflin	0.93	1.09	10137	0.40	0.00
Scott Foresman	1.88	0.80	10137	0.02	1.88
Success for All	1.84	1.50	10137	0.22	0.00
Higher-SES slope					
Harcourt	3.85	0.08	948	0.00	3.85
Open Court	−1.62	0.37	948	0.00	2.23
Reading Mastery	−0.31	0.29	948	0.28	3.85
Houghton Mifflin	−0.12	0.15	948	0.44	3.85
Scott Foresman	0.01	0.15	948	0.94	3.85
Success for All	−1.03	0.26	948	0.00	2.82
Adjustment from higher-SES slope for lower-SES					
Harcourt	0.19	0.12	10137	0.12	0.00
Open Court	0.43	0.45	10137	0.34	0.00
Reading Mastery	−0.42	0.35	10137	0.24	0.00
Houghton Mifflin	0.04	0.29	10137	0.88	0.00
Scott Foresman	0.29	0.24	10137	0.24	0.00
Success for All	−0.14	0.45	10137	0.76	0.00
Lower-SES on intercept	0.91	0.01	10137.00	0.00	
Lower-SES on slope	0.00	0.00	10137.00	0.01	

Random Effects	Variance	Chi square	*p* value
Class	34.91	2649.60	<0.01
Linear	89.60		
Child	127.02	24574.90	<0.01

Note. Calculated score denotes the adjusted mean or slope for each curriculum, depending on its significance from the reference group (Harcourt).

was ORF, which relies heavily on fluent and accurate decoding skills. It may be that results would have differed had the reading outcome been of comprehension or other higher order literacy skills.

Whereas score differences were statistically significant, effect sizes, which ranged from negligible to moderate, suggest some findings to be more practically important than others. When interpreting differences between curricula it is important to examine the practical significance of score differences. Nevertheless, small effects accumulating over time may become educationally important (Sanders & Horn, 1998; Xue & Meisels, 2004). It is also important to keep in mind that this is a descriptive and correlational study and that curricula were not randomly assigned to schools. Thus, bias associated with curriculum selection could exist. For example, schools may have selected a particular curriculum for a particular reason. Indeed, the results suggest that schools might have selected a curriculum specifically because they had many children beginning school with weak early reading skills. For example, in first grade, students whose schools selected Reading Mastery tended to have lower initial reading scores than students whose schools selected other curricula.

The third research question led us to investigate the possible differential impact of curricula for children from lower-SES homes compared to their more affluent peers. Results indicate that, as expected, differences in ORF existed between lower and higher-SES readers. Although some differences were small, albeit significant, it appears that certain curricula may be associated with higher ORF growth, especially for lower-SES students. This depended, however, on grade level. For example, in first and second grades, children from lower-SES homes achieved generally lower ORF scores regardless of curriculum. However, in second grade there were fewer differences and by third grade there was generally no difference in lower- and higher-SES students' performance across curricula, with the notable exception of Scott Foresman, for which children from lower-SES homes achieved higher ORF scores than did their higher-SES peers. Still, no SES difference was found to be greater than 7 CRW/min, with the difference being as few as 2 CRW/min in some cases. Once SES and initial status were held constant, differences between curricula remained but demonstrated a much smaller effect. Houghton Mifflin, one of the lowest scoring curricula at all grade levels, remained as such, with ORF achievement that was significantly lower than Harcourt. Relatively small differences suggest that, although lower-SES students appear to have slightly lower scores, the practical difference between curricula is not great. It is also important to consider that our sample was drawn from schools receiving Reading First funds. As such, the students in the study were generally less affluent than the student population as a whole. There may have been larger differences if a greater number of higher-SES schools had been included in this sample.

Limitations and Cautions

There are a number of limitations to this study, which should be considered when interpreting the results. First, these results provide only correlational indications of curriculum effects on student reading skill growth. Second, although Reading First site visits were designed to monitor implementation, and school reading coaches were assigned to managing and ensuring the use of the core, it is possible that other factors within and across schools could have contributed to our results. For example, it is possible that some curricula were used more effectively because teachers received extra professional development through Reading First. Schools may have used different editions of each curriculum, although curriculum adoptions by schools occurred during the 2004 school year, which makes this less likely. Teachers may have used other supplementary programs and materials alongside the

core curriculum. Because supplementary programs are not coded at the student level, it is impossible to examine whether these programs influenced growth of ORF. This limitation is moderated somewhat by the constraints of the Reading First program which required that schools follow the core curriculum in order to maintain funding. However, results may be due in part to supplementary materials and not solely the role of the core. Third, because only one measure of reading was used (ORF), the relation of curricula to other literacy or language skills could reveal different results than those presented here. Although the ORF measure was administered by trained staff who had been instructed on proper administration procedures, differences across sites could impact the reported results of the ORF assessments. In terms of benchmarks, those used for Reading First schools are based on norm-reference estimates of student performance and did not originate as a result of experimental analyses. Results might vary depending on the benchmarks used.

Implications

As teachers and instructional leaders make decisions about reading instruction and materials for purchase and use, particularly for students from lower-SES homes, these results suggest that a well-designed, evidence-based core curriculum can assist in raising and sustaining students' achievement. There were differences in curriculum effects on students' achievement, but this was complicated by students' SES status. However, the differences were relatively small, therefore making it difficult to judge their practical importance. Thus, the inclusion of specific curricula in policies and reform efforts should be tempered first by the systematic and scientific testing of instructional materials. This will help to ensure that materials that are recommended for schools will help them achieve the desired academic outcomes.

The differences observed in students' reading skill growth across curricula also begs the question, why isn't there better efficacy and effectiveness research on published core curricula? Presently, as in years past, school districts spend millions of dollars with no assurance that the purchased curriculum will serve the purpose for which it was designed—improving the reading skills of their students. As schools and teachers are held increasingly accountable for their students' outcomes, randomized control field trials which account for multiple student and school factors through the appropriate statistical treatment of the data would help school districts make informed decisions about how to use their limited resources. Moreover, as accumulating evidence suggests (Connor, Morrison, Fishman et al., 2007; Foorman, Francis, Fletcher, Schatschneider, & Mehta, 1998; Hamre & Pianta, 2005), adjusting instruction according to students' individual language and literacy needs may be more effective than more global "one size fits all" approaches. A strong evidence-based core curriculum in the hands of a well-trained teacher may provide a foundation for effective literacy instruction, but the curriculum alone may be insufficient for promoting effective instruction for all children.

References

Aikens, N. L., & Barbarin, O. (2008). Socioeconomic differences in reading trajectories: The contribution of family, neighborhood, and school contexts. *Journal of Educational Psychology, 100*(2), 235–251.

Arnold, D. H., & Doctoroff, G. L. (2003). The early education of socioeconomically disadvantaged children. *Annual Review of Psychology, 54*, 517–545.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A powerful and practical approach to multiple testing. *Journal of the Royal Statistical Society, Series B, 57*(1), 289–300.

Buck, J., & Torgesen, J. K. (2006). *The relationship between performance on a measure of oral reading fluency and performance on the Florida Comprehensive Assessment Test.* Retrieved August 3, 2006, from http://www.fcrr.org/TechnicalReports/TechnicalReport1.pdf

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* Erlbaum: Hillsdale, NJ.

Connor, C. M., Morrison, F. J., Fishman, B. J., Schatschneider, C., & Underwood, P. (2007). The early years: Algorithm-guided individualized reading instruction. *Science, 315*(5811), 464–465.

Connor, C. M., Piasta, S. B., Glasney, S., Schatschneider, C., Fishman, B., & Underwood, P. (2007). Individualizing student instruction precisely: Effects of child by instruction interactions on first graders' literacy development. Manuscript in review, Child Development.

Connor, C. M., Son, S., Hindman, A., & 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.

Crawford, L., Tindal, G., & Stiver, S. (2001). Using oral reading rate to predict student performance on statewide assessment tests. *Educational Assessment*, 7(4), 303–323.

Duncan, G. J., & Brooks-Gunn, J. (2000). Family poverty, welfare reform, and child development. *Child Development, 71*(1), 188–196.

Duncan, G. J., & Raudenbush, S. W. (1999). Assessing the effects of context in studies of child and youth development. *Educational Psychologist, 34*(1), 29–41.

Duncan, G. J., Yeung, W. J., Brooks-Gunn, J., & Smith, J. R. (1998). How much does poverty affect the life chances of children? *American Sociological Review, 63*(3), 406–423.

Entwisle, D., & Alexander, K. (1993). Entry into school: The beginning school transition and educational stratification in the United States. *Annual Review of Sociology, 19*, 401–423.

Evans, G. W. (2004). The environment of childhood poverty. *American Psychologist, 59*, 77–92.

FCRR (2002). *Success for All Report*. Retrieved August 2, 2007, from http://www.fcrr.org/fcrrreports/pdf/success_for_all_report.pdf

FCRR (2006). *Summary Table for FCRR Reports*. Retrieved August 30, 2007, from http://www.fcrr.org/FCRRReports/CReports.aspx?rep=core

Foorman, B. R., Francis, D. J., Fletcher, J. M., Schatschneider, C., & Mehta, P. (1998). The role of instruction in learning to read: Preventing reading failure in at risk children. *Journal of Educational Psychology, 90*, 37–55.

Foorman, B. R., Schatschneider, C., Eakin, M. N., Fletcher, J. M., Moats, L. C., & Francis, D. J. (2006). The impact of instructional practices in grades 1 and 2 on reading and spelling achievement in high poverty schools. *Contemporary Educational Psychology, 31*, 1–29.

Foorman, B. R., & Torgesen, J. K. (2001). Critical elements of classroom and small-group instruction promotes reading success in all children. *Learning Disabilities Research & Practice, 16*(4), 203–212.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., Walz, L., & Germann, G. (1993). Formative evaluation of academic progress: How much growth can we expect? *School Psychology Review, 22*, 27–48.

Good, R. H., & Kaminski, R. A. (Eds.). (2002). *Dynamic Indicators of Basic Early Literacy Skills*, 6th ed. Eugene, OR: Institute for Development of Educational Achievement.

Good, R. H., Simmons, D. C., & Kame'enui, E. J. (2001). The importance and decision-making utility of a continuum of fluency-based indicators of foundational reading skills for third-grade high-stakes outcomes. *Scientific Studies of Reading, 5*(3), 257–288.

Hamre, B. K., & Pianta, R. C. (2005). Can instructional and emotional support in the first-grade classroom make a difference for children at risk of school failure? *Child Development, 76*(5), 949–967.

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children.* Baltimore: Paul H. Brookes Publishing.

Hasbrouck, J. E., & Tindal, G. (1992, Spring). Curriculum-based oral reading fluency norms for students in grades 2 through 5. *Teaching Exceptional Children*, 41–44.

Haycock, K. (2000). No more settling for less. *Thinking K-16, 4*, 3–16.

Juel, C. (1988). Learning to read and write: a longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology, 80*(4), 43–447.

Katz, L., Kling, J., & Liebman, J. (2003). Boston site findings: The early impacts of moving to opportunity. In John Goering & Judith D. Feins (Eds.), *Choosing a Better Life? Evaluating the Moving to Opportunity Social Experiment* (pp. 177–211). Washington, DC: Urban Institute Press.

Kozol, J. (1991). *Savage inequalities: Children in America's schools*. New York: Harper Perennial.

Ladd, H. F., & Ludwig, J. (2003). The Effects of MTO on educational opportunities in Baltimore. In John Goering & Judith D. Feins (Eds.), *Choosing a Better Life? Evaluating the Moving to Opportunity Social Experiment* (pp. 177–51). Washington, DC: Urban Institute Press.

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

Lee, J., Griggs, W. S., & Donahue, P. L. (2007). *The nation's report card: Reading 2007*. Washington, DC: National Center for Education Statistics.

McCoach, D. B., O'Connoll, A. A., Reis, S. M., & Levitt, H. A. (2006). Growing readers: A hierarchical linear model of children's reading growth during the first 2 years of school. *Journal of Educational Psychology, 98*(1), 14–28.

Miller, C. (2007). Reading First Update. Available at: http://www.justreadflorida.com

NICHD-ECCRN (2002). Child-care structure–process–outcome: Direct and indirect effects of child-care quality. *Psychological Science, 13*(2), 199–206.

NRP. (2000). National Reading Panel report: Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction (No. NIH Pub. No. 00-4769). Washington DC: U.S. Department of Health and Human Services, Public Health Service, National Institutes of Health, National Institute of Child Health and Human Development.

Pressley, M., Wharton-McDonald, R., Allington, R., Block, C. C., Morrow, L., Tracey, D., et al. (2001). A study of effective first-grade instruction. *Scientific Study of Reading, 5*, 35–58.

Pressley, M., Wharton-McDonald, R., Mistretta-Hapston, J., & Echevarria, M. (1998). Literacy instruction in 10 fourth and fifth grade classrooms in upstate New York. *Scientific Studies of Reading, 2*(2), 159–194.

Raudenbush, S. W. (2004). *Schooling, statistics, and poverty: Can we measure school improvement?* Educational Testing Services.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.) Thousand Oaks, CA: Sage Publications.

Rothstein, R. (2004). *Class and schools: Using social, economic, and educational reform to close the Black–White achievement gap*. Washington, DC: Economic Policy Institute.

Sanders, W. L., & Horn, S. P. (1998). Research findings from the Tennessee value-added assessment system (TVAAS) database: Implications for educational evaluation and research. *Journal of Personnel Evaluation in Education, 12*(3), 247–256.

Schatschneider, C., Buck, J., Torgesen, J. K., Wagner, R., Hassler, L., Hecht, S. A., et al. (2004). *A multivariate study of individual differences in performance on the reading portion of the Florida Comprehensive Assessment Test. A brief report*. Retrieved August 4, 2006, from http://www.fcrr.org/TechnicalReports/Multi_variate_study_december2004.pdf

Shaw, R., & Shaw, D. (2002*). DIBELS Oral Reading Fluency-Based Indicators of Third Grade Reading Skills for Colorado State Assessment Program (CSAP) (Technical Report)*. Eugene, OR: University of Oregon.

Shinn, M. R., Good, R. H., III, Knutson, N., & Tilly, W. D., III (1992). Curriculum-based measurement of oral reading fluency: A confirmatory analysis of its relation to reading. *School Psychology Review, 21*, 459–479.

Snow, C. E., Burns, M. S., & Griffin, P. (Eds.). (1998*). Preventing reading difficulties in young children*. Washington, DC: National Academy Press.

Spira, E. G., Bracken, S. S., & Fischel, J. E. (2005). Predicting improvement after first-grade reading difficulties: The effects of oral language, emergent literacy, and behavior skills. *Developmental Psychology, 41*(1), 225–234.

Taylor, B. M., Pearson, D. P., Clark, K., & Walpole, S. (2000). Effective schools and accomplished teachers: lessons about primary-grade reading instruction in low-income schools. *The Elementary School Journal, 101*(2), 121–165.

Tindal, G., & Marston, D. (1996). Reflections on technical adequacy of alternative reading measures as performance assessments. *Exceptionality, 6*(4), 247–251.

Torgesen, J. K., Alexander, A. W., Wagner, R. K., Rashotte, C. A., Voelier, K. K. S., & Conway, T. (2001). Intensive remedial instruction for children with severe reading disabilities: Immediate and long-term outcomes from two instructional approaches. *Journal of Learning Disabilities, 34*(1), 33–58.

U.S. Department of Education (2007). *Reading First Implementation Evaluation: Interim Report*. Washington, DC: Office of Planning, Evaluation, and Policy Development, Policy and Program Studies Service.

Vernon-Feagans, L., Hammer, C. S., Miccio, A., & Manlove, E. (2002). Early language and literacy skills in low-income African American and Hispanic children. In S. B. Neuman & D. K. Dickinson (Eds.), *Handbook of early literacy research* (pp. 192–210). New York: Guilford Press.

Wharton-McDonald, R., Pressley, M., & Hampston, J. M. (1998). Literacy instruction in nine first-grade classrooms: Teacher characteristics and student achievement. *Elementary School Journal, 99*(2), 101–128.

Whitehurst, G. J., & Lonigan, C. J. (1998). Child development and emergent literacy. *Child Development, 69*(3), 848–872.

Wyner, J. S., Bridgeland, J. M., & Diiulio, J. J. (2007). *Achievement rap: How America is failing millions of high-achieving students from low- income families.* Jack Kent

Cooke Foundation. Available at http://www.civicenterprises.net/pdfs/jkc.pdf

Xue, Y., & Meisels, S. J. (2004). Early literacy instruction and learning in kindergarten: Evidence from the early childhood longitudinal study—Kindergarten class of 1998–1999. *American Educational Research Journal*, *41*(1), 191–229.

Yeung, W. J., Linver, M. R., & Brooks-Gunn, J. (2002). How money matters for young children's development: Parental investment and family processes. *Child Development*, *73*, 1861–1879.