Read Writ (2014) 27:451–479
DOI 10.1007/s11145-013-9452-5

Predictors of at-risk kindergarteners' later reading difficulty: examining learner-by-intervention interactions

Deborah C. Simmons · Aaron B. Taylor · Eric L. Oslund · Leslie E. Simmons · Michael D. Coyne · Mary E. Little · D'Ann M. Rawlinson · Shanna Hagan-Burke · Oi-man Kwok · Minjung Kim

Published online: 22 May 2013

Abstract This longitudinal study examined (a) the second-grade reading outcomes of 368 children who participated in either experimental or school-designed supplemental intervention in kindergarten, and (b) the influence and interactions of learner variables and type of intervention on reading achievement. Descriptive findings indicated that percentages of students identified as at risk or not at risk were statistically comparable between interventions. Entry-level letter identification scores predicted performance on all second-grade reading outcomes for students in both interventions. The influence of entry-level sound matching, receptive vocabulary knowledge, English language learner status and reading outcomes was moderated by type of intervention. Performance on a curriculum-embedded measure administered 8 weeks into intervention reliably predicted second grade reading performance.

Keywords Predictors of reading achievement · Early reading development

D. C. Simmons (✉) · L. E. Simmons · S. Hagan-Burke · O. Kwok · M. Kim
Department of Educational Psychology, Texas A&M University, 4225 TAMU, 704 Harrington Tower, College Station, TX 77843-4225, USA
e-mail: dsimmons@tamu.edu

L. E. Simmons
e-mail: lsimmons@tamu.edu

S. Hagan-Burke
e-mail: shaganburke@tamu.edu

O. Kwok
e-mail: Omkwok@neo.tamu.edu

M. Kim
e-mail: mjkim@neo.tamu.edu

A. B. Taylor
Department of Psychology, Texas A&M University, 4235 TAMU, College Station, TX 77843-4235, USA
e-mail: aaron.taylor@tamu.edu

Introduction

Kindergarten represents a critical window of opportunity in which intervention can differentially accelerate reading growth compared to later intervention for children at early reading risk (Al Otaiba & Fuchs, 2006; Simmons et al., 2008; Vadasy, Sanders, & Peyton, 2006; Vellutino, Scanlon, Small, & Fanuele, 2006). Converging evidence has shown the benefit of explicit, systematic intervention in phonemic awareness and decoding for children identified as being at risk of reading difficulty in the early grades (Cavanaugh, Kim, Wanzek, & Vaughn, 2004; Foorman, Breier, & Fletcher, 2003; Scammaca, Vaughn, Roberts, Wanzek, & Torgesen, 2007). Across these studies, two prominent findings emerged. First, the majority of children attained adequate levels of achievement in reading-related skills (e.g., letter-naming, phonemic awareness) and early word reading outcomes at the end of kindergarten intervention. Second, a small but significant percentage of children failed to benefit adequately from supplemental or additional small-group kindergarten intervention, requiring more intensive and extensive instructional support (e.g., Vellutino et al., 2006).

A critical assumption of early intervention research is that benefits maintain over time. A second assumption is that through retrospective analyses of learner characteristics we can more effectively identify kindergarteners who will not attain adequate benefit from supplemental intervention. Empirical research into variables that predict later reading achievement of students who participated in small-group, supplemental kindergarten intervention is emerging, but to date, it is based on a relatively small number of experimental studies with varying methodologies (Scammaca et al., 2007). Unlike prior research that has examined factors predicting student achievement post intervention, this study examined whether type of intervention moderated the influence of predictors that have been investigated in contemporary research.

To address this issue, we examined the reading performance of children 2 years post kindergarten reading intervention, focusing on the relationship between kindergarteners' entry-level reading-related, demographic, and cognitive processing

E. L. Oslund
Meadows Center for Preventing Educational Risk, The University of Texas at Austin, 1912 Speedway, D4900, Austin, TX 78712, USA
e-mail: ericoslund@yahoo.com

M. D. Coyne
Department of Educational Psychology, Neag School of Education, University of Connecticut, 249 Glenbrook Road, Unit 3064, Storrs, CT 06269-3064, USA
e-mail: mike.coyne@uconn.edu

M. E. Little
Department of Child, Family and Community Sciences, University of Central Florida, P.O. Box 161250, Orlando, FL 32816-1250, USA
e-mail: mary.little@ucf.edu

D. M. Rawlinson
University of Central Florida, 4000 Central Florida Boulevard TA 93 Room 403, Orlando, FL 32816, USA
e-mail: D'Ann.Rawlinson@ucf.edu

performance on end-of-second-grade outcomes and their interaction with type of intervention (explicit, code-based experimental and school-designed comparison). We also conducted secondary data analyses involving the experimental condition only to investigate whether student performance early in the intervention process was associated with later reading achievement.

We conducted the current study to extend our understanding of later reading achievement by children who participate in supplemental kindergarten reading intervention, and factors that predict reading outcomes 2 years post intervention. We define supplemental intervention as instruction that occurs in small groups in addition to core instruction that all students typically receive. The data reported in this article are combined from two previous randomized-controlled investigations that compared the efficacy of two supplemental reading interventions, a code-based commercial program (early reading intervention; Pearson/Scott Foresman, 2004) and school-designed reading intervention, on kindergarten students' reading achievement in three states (Texas, Connecticut, and Florida). In the intervention studies (Coyne et al., 2013; Simmons et al., 2011), we investigated end-of-kindergarten effects on the reading achievement of children identified as at risk of reading difficulties who were randomly assigned by classroom to either the experimental early reading intervention (ERI) or the school-designed intervention (SDI).

Conceptual framework for examining predictors of later reading achievement

To advance our understanding of factors associated with later reading achievement, our conceptual framework connected (a) empirical research on learner and intervention characteristics that are associated with children's response to intervention, and (b) emerging research on child-by-instruction interaction effects on reading. We selected learner variables based on prior predictor studies (e.g. Schatschneider, Fletcher, Francis, Carlson, & Foorman, 2004) and previous intervention followup studies (e.g., Al Otaiba & Fuchs, 2006) and examined their predictive validity by intervention. For the experimental intervention only, we augmented the predictor variable set with an early indicator of performance administered after 8 weeks of intervention and examined its association with later achievement. We hypothesized that a child's early response to intervention would be an important index of later reading achievement. As context for this study, we review prior research examining the long-term influence of kindergarten reading intervention on subsequent reading outcomes.

Longitudinal research on effects of supplemental kindergarten reading intervention

Supplemental reading intervention refers to instruction that is "layered on top of classroom reading instruction" (Vadasy, Sanders, & Abbott, 2008, p. 52) and provides students an additional opportunity to learn. Based on the accumulating empirical knowledge base in early reading intervention, schools are providing more systematic supplemental reading support in the primary grades, often as part of a response to intervention process (Gersten et al., 2009). The long-term effects of

supplemental kindergarten intervention have been evaluated through multiple methods (e.g., statistical comparison to control groups, absolute performance attainment [benchmarks], discrepancy from peers). For example, in their study of responsiveness to varying combinations of kindergarten and first-grade intervention, Al Otaiba and Fuchs (2006) classified children's performance at the end of first grade as nonresponsive, sometimes responsive, or always responsive. In kindergarten, nonresponsiveness was gauged as performing in the bottom 30 % of intervention participants, relative to growth on letter-sound and segmentation measures. In first grade, students were considered responsive if their oral reading fluency score was at or above the intervention group's mean. Across kindergarten and first grade, 53 % of students who participated in the kindergarten intervention only were classified as always responsive, 35 % as sometimes responsive, and 13 % as nonresponsive. At the end of third grade, none of the students identified as responsive or sometimes responsive was reported by their schools as having reading problems.

In a followup study of 41 children who received kindergarten intervention, (Simmons et al., 2008) found that at the end of third grade, the majority responded strongly and maintained reading trajectories that exceeded the 50th percentile on most measures. In addition, reading risk indexed by performing at or below the 30th percentile on a range of standardized measures, was stable over time, with the exception of oral reading fluency, which evidenced greater variability. On measures of word reading, 88–94 % of students in the most effective intervention were no longer at risk at the beginning of first grade and remained out of risk through third grade. However, the number of children who remained at risk on the oral reading fluency measure across intervention groups was markedly higher, 38–51 %.

Vellutino et al. (2006) provided small-group, supplemental intervention to 117 children in kindergarten; children performing in the bottom half of participants also received supplemental intervention in first-grade. Remediation difficulty was benchmarked against standardized performance levels on third-grade assessments. Students who achieved average scores (standard score of 90 or above) were considered less difficult to remediate than students who continued to perform below 90 (i.e., below 25th percentile on nationally normed test) despite intervention. On third-grade assessments, 62 % of students who received only kindergarten intervention attained average-level reading; 16 % were classified as difficult to remediate despite kindergarten and first-grade intervention. The average performance of the difficult-to-remediate group hovered slightly above the 30th percentile on word identification and word attack at the end of first grade, evidencing downward trends in second and third grade. At the end of third grade, average performance on word reading skills fell to the 8–9th percentile.

The influence and importance of benchmarks and measures in determining response to intervention was documented in a secondary data analysis of reading achievement from students who participated in intensive first-grade intervention (Barth et al., 2008). Response profiles were based on four measures that assessed a range of word reading, oral reading fluency, and comprehension skills, and used a cutpoint of 0.5 standard deviations below the national norm to index response. Researchers found that responders were more reliably identified than nonresponders;

nonetheless, response profiles across the four measures ranged from 0.003 to 0.50. When all four measures were used, 50 % of the participants were identified as nonresponders.

Findings from this set of studies consistently indicate that kindergarten intervention provides an important jumpstart for the majority of students who receive it. Across these studies, nonetheless, the measures, benchmark criteria, and reference group used to determine responsiveness or risk differed. Importantly, in all the studies a percentage of students remained at risk of later reading difficulty.

Learner and intervention factors associated with response to kindergarten intervention

Empirically derived predictors of response to intervention

Schatschneider et al. (2004) acknowledged that identifying which variables can be used to predict reading achievement prior to the onset of reading instruction is an age old problem. A converging body of research has identified predictors of reading achievement among the range of readers in entire classrooms (Al Otaiba et al., 2011; Hogan & Thomson, 2010; Schatschneider et al., 2004). Across these studies, measures of letter knowledge, phonological awareness, and naming speed consistently predicted reading achievement in the primary grades. To better understand the factors associated with children's response to intervention, researchers have conducted retrospective analyses and syntheses of intervention research (e.g., Al Otaiba & Fuchs, 2002; Chard et al., 2008; Coyne, Kame'enui, Simmons, & Harn, 2004; Nelson, Benner, & Gonzalez, 2003). The ultimate goal of these studies has been to identify the characteristics of children who are unlikely to attain adequate levels of reading even with current interventions in an attempt to proactively design and deliver interventions that are more likely to address the individual needs of this small but significant group of learners for whom early intervention was not effective.

Across intervention studies, researchers commonly focus on student demographics, cognitive processing, phonological processing, and entry-level reading-related skills as predictors of reading outcomes. Findings of studies including kindergarten students indicate a common set of student-related factors that are reliable predictors of reading performance. Specifically, individual differences in phonemic awareness, alphabetic knowledge, and rapid automatized naming (RAN) in kindergarten have consistently been found to predict reading achievement (Al Otaiba & Fuchs, 2002; Nelson et al., 2003; Vellutino et al., 1996). Vellutino et al. (2006) documented that kindergarten children with significant difficulties in cognitive abilities that are foundational to reading ability (e.g., phonological awareness, letter identification, retrieval speed) may experience sustained reading difficulties despite early intervention. In a followup to their intervention study, Al Otaiba and Fuchs (2006) found that a combination of student characteristics measured post intervention (i.e., vocabulary, rapid naming, problem behavior, verbal memory, syntactic awareness) and years of intervention reliably classified 82 % of responsive

and 84 % of nonresponsive students. Across these studies, researchers have consistently documented linkages between student performance on cognitive and reading-related measures and reading outcomes.

Early response to intervention

Across intervention studies, authors have commonly found that a child's early response to intervention is a reliable predictor of later reading achievement (e.g., Al Otaiba & Fuchs, 2006; Simmons et al., 2008). Vellutino et al. (2006) concluded that a child's initial response to intervention as gauged by end of kindergarten performance can be a "good barometer" (p. 166) of how readily skills can be remediated. While end of kindergarten may reliably predict later reading achievement, the fundamental problem is that once reading trajectories are established, they are difficult to alter. Given the results of prevention studies indicating that the majority of students in the bottom 25 % in kindergarten can attain adequate levels of reading proficiency (Foorman et al., 2003; Torgesen, 2000), a critical issue is how to proactively identify students and intensify intervention appropriately. Thus, for students who remain at risk of reading difficulty at the end of kindergarten even after supplemental reading intervention, we have lost a critical window of opportunity to provide intensive intervention.

Relatively few intervention studies have examined the influence of the variables prospectively; that is, collected variables prior to intervention (Vellutino et al., 2006). While retrospective analyses from intervention studies have identified variables that predict kindergarteners' differential response to intervention, in some cases, these variables were collected post intervention (Al Otaiba & Fuchs, 2006). Moreover, few studies have included measures that index a child's performance early in the intervention process. A question of interest in followup studies of kindergarten intervention is whether measures gathered prior to or in the early stages of intervention predict achievement trajectories and, more important, can enable educators to differentiate intervention more effectively. We were interested in whether children's performance on assessments administered early in the intervention process would provide information that enabled practitioners to make more reliable decisions regarding children who may experience sustained reading difficulty.

Learner-by-intervention interactions

Considerable consistency has been established in the nature of learner skills and characteristics that predict later reading achievement. However, to date, only limited research has addressed the interaction of learner characteristics and type of supplemental intervention on the reading outcomes of children at risk of reading difficulties. Among the emerging body of research that has investigated learner-by-instructional practice interactions, the primary focus has been on skill differences among the range of learners in classrooms (e.g., low decoding skill versus high decoding skill; e.g., Connor, Morrison, & Katch, 2004; Foorman, Francis, Fletcher, Schatschneider, & Mehta, 1998).

Specifically, in a study of first-grade reading instruction in general education classes, Connor et al. (2004) found that types of intervention differentially predicted children's decoding outcomes. That is, children with initially low decoding skills benefited more from greater amounts of teacher-managed explicit decoding instruction than from child-managed instruction. By comparison, children with higher initial vocabulary scores benefited from greater amounts of child-managed activities opposed to teacher-directed instruction. In another study of first-grade instruction, Connor et al. (2009) concluded that predicting types and amounts of instruction based on student profiles is not a straightforward process. Specifically, these researchers found that instructional recommendations are complex and must be dynamic. Their findings indicated that precise amounts of individualized instruction predicted students' reading outcomes. Neither teacher qualifications nor school characteristics predicted student outcomes. Based on their findings, Connor et al. (2009) concluded that child-by-instruction interactions are causally related to student achievement. The majority of studies examining child-by-instruction interactions have been conducted, however, in general education classrooms involving the full range of learners. To our knowledge, learner-by-instruction interactions have not been examined for students participating in supplemental interventions.

The present study

In the present study, we examined the end-of-second-grade reading performance of kindergarteners who had participated in supplemental reading intervention and investigated learner and intervention variables that influenced those outcomes. Our research sought to extend current literature in four critical ways. First, we examined reading performance 2 years post intervention on important transfer skills, including word identification, oral reading fluency, and passage comprehension (Compton et al., 2005). Descriptive analyses were conducted to examine students' absolute performance relative to established benchmarks (i.e., 25th percentile). Second, we sought to increase the reliability of identifying students' risk status (at risk; not at risk). That is, rather than establishing a single cut point and dichotomizing students into the two categories, which is almost certain to misclassify at least some students given measurement error, we added a statistical buffer of 1.96 standard errors of measurement in our classification. Third, we included predictor variables (i.e., entry-level reading-related skills and student demographics) collected prior to intervention and examined the interaction of predictors and intervention on outcomes. Finally, we conducted secondary analyses in the experimental intervention and expanded the predictor set to include a measure of early response to kindergarten intervention (i.e., performance on a curriculum-embedded test). We investigated the three following questions:

1. What performance levels do students who participated in supplemental kindergarten reading intervention attain at the end of second grade? What percentages of students remain at risk of reading difficulty?
2. What is the relationship between kindergarteners' demographics, entry-level reading-related skills, and type of intervention on end-of-second-grade reading

outcomes? Is the influence of predictors moderated by type of intervention (experimental vs. school-designed)?
3. Does a measure of early response (i.e., curriculum-embedded measure) reliably predict later reading risk?

Methods

Research context

This study extended research from two previous investigations that compared the effects of supplemental reading interventions on the performance of kindergarten students identified as being at risk of reading difficulty (Coyne et al., 2013; Simmons et al., 2011). In the earlier studies, students were assigned randomly at the classroom level to either to an explicit/systematic commercial program (experimental group) or a school-designed typical practice intervention (comparison group). The first study was conducted in 2006–2007 in Texas and Connecticut; the second was conducted in 2007–2008 in Florida. Across the studies and interventions, school-based interventionists that included general education teachers, reading specialists, and paraprofessionals were randomly assigned to provide experimental or school-designed intervention.

To increase comparability between interventions and between studies, a number of common instructional components were standardized. Groups in both interventions consisted of 3–5 students. Interventionists in both interventions were asked to meet with their groups for 30 min, 5 days per week between pretest and posttest (approximately 100 lessons). Schools coordinated intervention delivery times according to their respective schedules; all intervention supplements occurred during the regular school day. Content in both interventions focused on early literacy skills, with experimental interventionists implementing the ERI (Pearson/Scott Foresman, 2004) curriculum and comparison interventionists providing school-designed reading intervention. Findings from the original randomized control trial indicated statistically and practically significant differences favoring the experimental condition on measures of phonemic awareness, letter names and sounds, and word attack. No statistically significant between-group differences were detected on measures of word identification, spelling, or passage comprehension. In the replication study conducted in Florida, no statistically significant differences between the ERI and SDI conditions were found on any posttest measure.

Interventions

Experimental intervention

Students in the experimental intervention in both studies received the ERI program, which includes 126 daily lessons of explicit, systematic instruction. Specifically, the program introduces phonemic, alphabetic, word reading, and spelling skills in three-

day cycles in which teachers model the skill and provide students multiple opportunities to respond and receive feedback. A typical ERI consisted of seven activities, each designed to last 3–5 min. The first 15 min of the lesson focuses on phonological awareness and alphabetic understanding; the second 15 min integrates writing and spelling with previously taught phonemic and alphabetic skills. The program is organized into four parts that progress from letters and sounds to word reading followed by sentence and storybook reading. The 126 lessons are scripted, providing specific language for teachers to ensure consistent communication of information. Each lesson that introduces new information includes a specified number of instructional interactions in which the teacher first models the information using consistent wording. Students practice the new skill with the teacher and then apply it to discrimination or generalization tasks. In addition, skills are carefully integrated to enhance learning. Scheduled instruction, review, and feedback are explicitly incorporated into the program. As delivered, this intervention would be classified as a standard protocol implementation.

School-designed interventions

The content of school-designed interventions (SDI) instructional approaches were allowed to vary naturally by school and interventionist. Although some schools coordinated the intervention at a school level, most allowed interventionists to design and deliver intervention individually. Teachers used a range of materials, including commercial materials exclusively, self-developed lessons and materials, and a combination of commercial and teacher-developed materials.

Interventionists received no professional development as part of the study; however, schools were not prohibited from providing professional development. Observations of school-designed interventions revealed that interventionists in both studies allocated time to a range of skills, including phonemic awareness, word reading, spelling, reading connected text, vocabulary, and listening comprehension. Teachers were observed, and their quality of instruction evaluated on a 4-point scale (1 = low; 4 = high). Mean quality scores for SDI interventionists across studies and interventions ranged from 2.55 to 3.02, indicating that the quality of instruction provided to students was fair to good.

Participants

Participants were children from 22 schools in three states (Connecticut, Florida, and Texas). As a criterion for involvement, participating schools agreed to provide supplemental reading instruction for children with identified reading risk. During the third week of kindergarten, researchers consulted with school personnel to identify children who (a) were considered in need of supplemental reading instruction, (b) were at least 5 years of age, and (c) received reading instruction in English.

Children qualified to participate in the study by being among the five lowest performers on screening measures within their class and meeting one or both of the following criteria: (a) a score at or below the 33rd percentile on the Dynamic

Indicators of Basic Early Literacy Skills (DIBELS; Good & Kaminski, 2002) Letter-Naming Fluency (LNF) measure (i.e., fewer than 6 letters correctly named in 1 min); or (b) a score below the 37th percentile on the Comprehensive Test of Phonological Processing (CTOPP; Wagner, Torgesen, & Rashotte, 1999) Sound-Matching (SM) subtest.

Across settings and years, 368 children participated in interventions (199 in ERI and 169 in SDI) and were present for at least one outcome measurement. Baseline demographic and achievement information are reported in Table 1. There were no statistically significant between-group differences on any demographic variable. Of the full sample of children who participated in intervention, 47 % were girls and 53 % were boys. The mean age of the sample at pretest was 5.49 years. The percentage of English language learners was comparable between groups (23 % SDI; 20 % ERI) as were the percentages of students eligible to receive speech/language services (11.9 % SDI; 13 % ERI). From the end of kindergarten through the end of second grade, the attrition total was 115 students, or 31 % of the original sample. Each followup sample was compared to the original sample using Chi square tests. Findings indicated no statistically significant between-sample differences on any demographic variable.

Measures and assessment procedures

Outcome measures

Oral reading fluency scores gathered at the beginning and end of first and second grade were included in the analyses. The word identification scores used were gathered at the end of kindergarten, the beginning and end of first grade, and the beginning and end of second grade. Passage comprehension scores gathered at the end of kindergarten and the end of first grade and the beginning and end of second grade were included in the analyses.

Oral reading fluency First- and second-grade measures used the Oral Reading Fluency (ORF) test from DIBELS (Good & Kaminski, 2002), which measures the accuracy and rate with which a student can read connected text. Students read a passage aloud for 1 min, and scores are based on the number of words correctly read. There are 20 alternate forms; alternate-form reliabilities range from 0.89 to 0.96. Concurrent validity based on the regional sample of students range from 0.91 to 0.96 as reported in DIBELS technical reports (University of Oregon Center on Teaching and Learning, 2002).

Word identification The Word Identification (WI) subtest from the Woodcock Reading Mastery Tests-Revised/Normative Update (WRMT-R/NU; Woodcock, 1987, 1998) measures a student's ability to read words in isolation. When taking the test, the student is presented with a written word and asked to identify the word with the correct pronunciation within 5 s. The 106 items increase in difficulty. The median split-half reliability is 0.97. Concurrent validity for first grade ranges from 0.48 to 0.82; it is not reported for either kindergarten or second grade.

Table 1 Demographic variables and predictor measure score means and standard deviations

Variable	SDI (N = 169)		ERI (N = 199)	
	N	(%)	N	(%)
Gender				
Male	85	50.30	90	45.24
Female	84	49.70	109	54.77
Ethnicity				
Asian	0	0.00	2	1.00
American Indian or Alaska Native	0	0.00	1	0.50
Black or African-American	25	14.79	29	14.57
Hispanic or Latino	58	34.31	67	33.66
White	83	49.11	94	47.24
Other	3	1.78	6	3.01
Speech or language eligible	20	11.83	22	11.01
English language learner	39	23.07	39	19.60

Variable	Mean (SD)	Mean (SD)
Age	5.47 (0.30)	5.51 (0.33)
Letter ID	88.99 (10.93)	88.30 (10.25)
Sound matching	7.71 (1.19)	7.64 (1.10)
Rapid object naming	8.37 (2.79)	8.17 (2.96)
PPVT	90.39 (13.98)	88.86 (14.40)

The *n* for School Designed Intervention condition was 114 and *n* for the ERI condition was 139 at the end of second grade. Chi square tests indicated no statistically significant differences between groups on any demographic variables at the end of second grade
SDI school-designed intervention; *ERI* early reading intervention

Passage comprehension In the Passage Comprehension (PC) subtest from the WRMT-R/NU (Woodcock, 1987, 1998), students are presented with passages between one and three sentences long from which a word is missing. They are then asked to identify the missing word that would correctly fit in the passage. For roughly one-third of the passages, an accompanying picture is related to the passage, enabling assessment of younger students. The median split-half reliability is 0.92. Concurrent validity in first grade ranges from 0.39 to 0.71; it is not reported for kindergarten or second grade.

Predictor measures

Predictor measures were based on previous studies of student demographics, reading-related measures, and quantity and quality of instruction associated with student responsiveness to intervention (Al Otaiba & Fuchs, 2006; Nelson et al., 2003; Vellutino et al., 2006). Pre-intervention reading-related measures, including rapid naming, phonemic sound matching, letter name identification (hereafter referred to as letter identification), and receptive vocabulary, were administered

prior to intervention in the fall of kindergarten. Student response to intervention was assessed through a curriculum-embedded test administered at the end of Part I of the program (i.e., after lesson 42) after 8 weeks of intervention.

Letter identification The Letter Identification (LI) subtest from the WRMT-R/NU (Woodcock, 1987, 1998) is a standardized, untimed test battery that assesses a student's ability to name upper- and lower-case letters presented in various fonts. The median split-half reliability for this measure is 0.84; no concurrent validity is reported for kindergarten.

Sound matching The Sound-Matching (SM) subtest of the CTOPP (Wagner et al., 1999) measures a child's ability to select one of three pictures that has the same initial or final sound in a target picture. This is an untimed measure. The internal consistency of this subtest, as measured by Cronbach's alpha, is 0.93 for both 5- and 6-year-old students. Item discrimination coefficients, which measure content validity, are 0.61 and 0.60 for 5- and 6-year-olds, respectively.

Rapid object naming The Rapid Object Naming (RON) subtest from the CTOPP (Wagner et al., 1999) measures the speed with which a student can name common objects. The student is presented with 72 pictures, and the score is the total number of seconds it takes to complete the task. Alternate-form reliability is 0.82 for 5-year-olds and 0.81 for 6-year-olds. The criterion-prediction validity is 0.37, as measured by the correlation between RON and the Total Word Reading Efficiency score from the Test of Word Reading Efficiency (TOWRE; Torgesen, Wagner, & Rashotte, 1999).

Vocabulary Vocabulary was assessed at pretest only, using the Peabody Picture Vocabulary Test—Third Edition (PPVT-III; Dunn & Dunn, 1997), an individually administered oral test of receptive vocabulary. For each test item, the student is presented with four black-and-white illustrations and asked to select the picture considered to best illustrate the meaning of a word presented orally by the examiner. The internal-consistency reliability (Cronbach's alpha) for form IIIA is 0.95 for the 5- to 6-year-old group, and the test–retest reliability ranges from 0.92 to 0.93. The concurrent validity with the Oral and Written Language Scales (OWLS; Carrow-Woolfolk, 1995) overall Oral Composite is 0.82.

Curriculum-embedded formative measure As an index of early response, we used the first part-test in the ERI curriculum. The items on the part-test are based on the first 42 lessons (Part 1) and measure the extent to which the student has mastered the material. Part-test 1 comprises three subtests composed of phonemic and alphabetic tasks. Letter Names and Sounds require the student to provide the correct letter name and corresponding sound for each of 11 letters taught in that part. First and Last Sounds in Words requires the student to produce the beginning and ending sounds in words presented both orally and with an accompanying picture. Finally, for the Letter-Sound subtest, the child is given tiles with the letters *d*, *f*, *l*, *m*, *p*, *r*, *s*,

or *t* printed on them. The examiner presents a page with six pictures and three empty boxes below the pictures. After the examiner points to a picture and reads it aloud, the student is asked to put the letter tiles for the first and last sounds into their corresponding (i.e., first and third) boxes. The score is a percentage of points earned.

Design and data-analysis procedures

This study is a secondary analysis of data collected as part of a larger randomized controlled trial in which students were randomly assigned to experimental or comparison conditions. Two sets of analyses were performed in this study. First, we examined the performance levels of students who participated in supplemental kindergarten reading intervention and the percentages that remained at risk of reading difficulty at the end of second grade. Second, we analyzed the relation between kindergarteners' demographics, entry-level reading-related skill and their interaction with type of intervention on end-of second-grade reading outcomes.

Identifying children who remain at risk of reading difficulty

To identify the proportion of students who remained at risk and their performance relative to normative data at the end of second grade, we classified students into performance groups (at risk or not at risk) based on their scores in the spring of each school year (i.e., from spring of kindergarten through spring of second grade). At each measurement occasion, students were classified into one of the two performance groups separately on each of the three outcome variables, with the exception that oral reading fluency was not measured in the spring of kindergarten.

The cutoff for the at-risk group was set at the 25th percentile using national norms reported by each test developer. This criterion was selected based on prior research (e.g., Vellutino et al., 2006) and to provide a meaningful indicator (e.g., lowest quartile of the distribution). Children who scored at or below the 25th percentile at a particular measurement occasion were classified in the at-risk group. The cutoff for the not-at-risk groups was different for the three outcome measures. For each, it was chosen to be 1.96 standard errors of measurement above the 25th percentile (standard errors of measurement were also taken from reports by the test developers), so that even the lowest performing children in the not-at-risk group could be considered significantly different from the highest performing children in the at-risk group (this assumes a normal distribution; 1.96 is the critical value for a *z* test).

In summary, by adding the statistically derived buffer between at-risk and not-at risk status, we sought to more accurately identify students who exceeded the 25th percentile and were no longer at risk of reading difficulty. This process yielded cutoff percentiles for not-at-risk groups at the 38th percentile for oral reading fluency, the 36th percentile for word identification, and the 46th percentile for passage comprehension. Although these percentiles provide a conservative estimate of students who were out of risk, we consider them an improvement over the more common approach of simply designating all students above the 25th percentile as out of risk because they take into account the variation (i.e., the standard error of the measurement) in measures we consider important for decision making.

Identifying predictors of risk

Ordinary least squares (OLS) regression was used to predict students' scores on each of the three outcome measures (oral reading fluency, word identification, passage comprehension) from their scores on the predictor measures and demographic characteristics. Scores on the outcome measures were taken at the end of second grade, but to maximize the sample available for analysis, and therefore increase the power to detect effects, for students who were not present at the end of second grade, their last available score was carried forward. To increase score interpretation and comparability, all scores were entered as percentiles based on test norms. All regression analyses were performed twice: once using the full sample with last available scores carried forward, and a second time using only the subsample of students present at the end of second grade.

Patterns of results were similar across the two sets of analyses—the median absolute difference in standardized regression coefficients comparing the two sets of analyses was 0.03—so only the full sample results are presented. The predictor measures included kindergarten scores on letter identification, sound matching, rapid object naming, and vocabulary (PPVT; Dunn & Dunn, 1997), as well as students' age, gender, speech-language eligibility, and ELL status (the last four predictors were entered using dummy codes). Amount of additional intervention and the state in which the student participated were also entered as control variables.

A separate regression model was estimated for each of the three outcome measures. Interactions between intervention and both the predictor measures and demographic characteristics were included to test whether the predictive ability of these variables differed by intervention. All continuous predictors were centered prior to being entered in interaction terms. Each model was also re-estimated with the curriculum-embedded formative measure included as a predictor. These models were estimated including only data from students in the ERI intervention, as the curriculum-embedded measure was only assessed in this intervention. Additionally, because the data for these models consisted of only one intervention, intervention and its interactions with demographic variables and predictor measures were dropped from the models. In all models, the predictor measures were entered as standard scores. To maintain the interpretability of the predictor variables' original scales and the dummy codes for some of the demographic variables, unstandardized rather than standardized coefficients are reported for the regression models.

Results

Performance levels using normative standards

Table 2 shows the percentages of students in each intervention in the two performance groups—at risk and not at risk (as well as the percentage in the middle group, labeled "marginal"). It is important to consider that outcomes were assessed through accuracy-based measures (word identification and passage comprehension) and fluency-based measures (oral reading fluency). While untimed accuracy

measures assess decoding proficiency, oral reading fluency assesses automaticity with word recognition, a skill that has been associated with reading comprehension (Johnson, Jenkins, Petscher, & Catts, 2009). For both interventions, the percentage of students in the not-at-risk group at their last measurement point varied by measure, with word identification having the highest percentage (74 %), followed by passage comprehension (60 %) and oral reading fluency (26 %). Some decline across time in the percentage of students categorized in the at-risk group was found for all three measures. By the end of second grade, oral reading fluency classified 55 % of students as at risk, word identification classified 10 %, and passage comprehension classified 15 %. While the smaller percentages for word identification and passage comprehension suggest students were performing above the 25th percentile, the larger percentage for oral reading fluency suggests they were not meeting expectations in this area. Chi square tests were used to test whether the distribution of students in the performance groups differed by intervention type at each measurement occasion for each outcome measure. None of the tests was significant, indicating that the distribution in the performance groups did not differ by condition.

To interpret performance relative to normative standards, mean percentile rankings for the full sample are shown in Table 3. Mean percentiles are reported to

Table 2 Percentages of students in performance groups by measurement occasion

Intervention	Spring kindergarten			Spring first grade			Spring second grade		
	At risk	Marginal	Not at risk	At risk	Marginal	Not at risk	At risk	Marginal	Not at risk
Oral reading fluency									
SDI	–[a]	–	–	63	10	27	56	20	24
ERI	–	–	–	63	9	28	54	18	28
Total	–	–	–	63	10	27	55	19	26
Word identification									
SDI	17	5	78	10	4	87	7	13	80
ERI	12	6	82	9	9	82	13	18	69
Total	14	5	80	9	7	84	10	16	74
Passage comprehension									
SDI	24	18	57	11	18	71	11	23	67
ERI	21	30	49	15	26	59	19	27	54
Total	22	25	53	13	22	64	15	25	60

Students were classified as at risk on a variable if they scored at or below the 25th percentile (using established norms, not sample values) at a particular measurement occasion. Students were classified as not at risk if they scored above 1.96 standard errors of measurement (using test developer-reported values) above the 25th percentile. The cutoffs were the 38th percentile for oral reading fluency, the 36th percentile for word identification, and the 46th percentile for passage comprehension. Students not classified as at risk or not at risk were classified as being at marginal risk. Chi square tests found no significant differences in the distribution of students into the three classes by intervention for any of the three variables at any of the three measurement occasions

[a] Oral reading fluency is not normed for children in kindergarten, so it was not measured in kindergarten

simplify findings as there were no statistically significant differences between groups. Because they are normed to the child's age, percentile scores do not show growth across time as clearly as do raw scores. Findings indicate that students classified as not at risk performed above the 60th percentile across all measures at the end of second grade. In contrast, the mean percentile rankings of students in the at risk group ranged from 9 on oral reading fluency to 16 on reading comprehension, corroborating the difficulties students experienced across outcome measures.

To display growth across time, plots of individual children's raw scores for each outcome measure across the five follow-up measurement occasions are shown in Fig. 1. Only a random sample of 100 children's trajectories is included to avoid over plotting. Mean trajectories for each performance group are plotted (solid lines) and, for comparison, raw scores corresponding to the normative 25th, 50th, and 75th percentiles are included (dashed lines).

Trend lines illustrate that oral reading fluency was the most challenging for the majority of students. On average, students in the at-risk group gained fewer than seven words correct per minute (CWPM) over the course of first grade (fall CWPM = 8.66 vs. spring CWPM = 15.10). In second grade, their CWPM growth averaged 23 words. Despite the second-grade growth, the mean ORF performance of children in the at-risk group was 41 CWPM, with an average percentile rank of 9.38 (see Table 3) at the end of second grade. While the ORF is not typically administered in the fall of first grade, our findings indicated that students in the not-at-risk group were reading 23 CWPM in the fall. Their average growth in first grade

Table 3 Normative percentile means and standard deviations by group across measurement occasions, full sample

Group	Measurement occasion					
	n	Kindergarten	First grade		Second grade	
		Spring *M* (SD)	Fall *M* (SD)	Spring *M* (SD)	Fall *M* (SD)	Spring *M* (SD)
Oral reading fluency						
At risk	174	–	15.81 (10.78)	10.16 (12.81)	14.00 (14.31)	9.38 (7.37)
Marginal risk	57	–	27.16 (13.36)	28.32 (18.47)	41.64 (17.43)	31.63 (3.23)
Not at risk	83	–	39.61 (19.63)	55.09 (22.81)	64.82 (17.42)	60.23 (14.15)
Word identification						
At risk	43	22.79 (19.84)	27.50 (18.23)	20.52 (16.64)	16.25 (12.69)	15.38 (6.74)
Marginal risk	48	47.67 (22.34)	45.65 (18.48)	43.64 (17.68)	30.40 (10.87)	31.20 (3.60)
Not at risk	277	71.52 (20.51)	72.69 (18.48)	77.12 (16.74)	65.98 (18.21)	66.66 (16.38)
Passage comprehension						
At risk	54	27.71 (20.43)	–	24.59 (17.05)	16.08 (11.31)	16.13 (8.13)
Marginal risk	85	38.89 (16.93)	–	40.84 (14.01)	36.73 (12.03)	35.97 (6.29)
Not at risk	199	58.36 (20.00)	–	67.81 (15.81)	65.07 (16.32)	67.17 (12.73)

Normative percentiles are not available for oral reading fluency prior to the fall of first grade. Passage comprehension was not measured in the fall of first grade

was 39 CWPM and growth in second grade was 48 CWPM, with an average ending score of 111 CWPM (60th percentile). Growth in word identification and passage comprehension was more consistent across time, with word identification appearing to be the easiest measure.

Predictors of end-of-second-grade risk

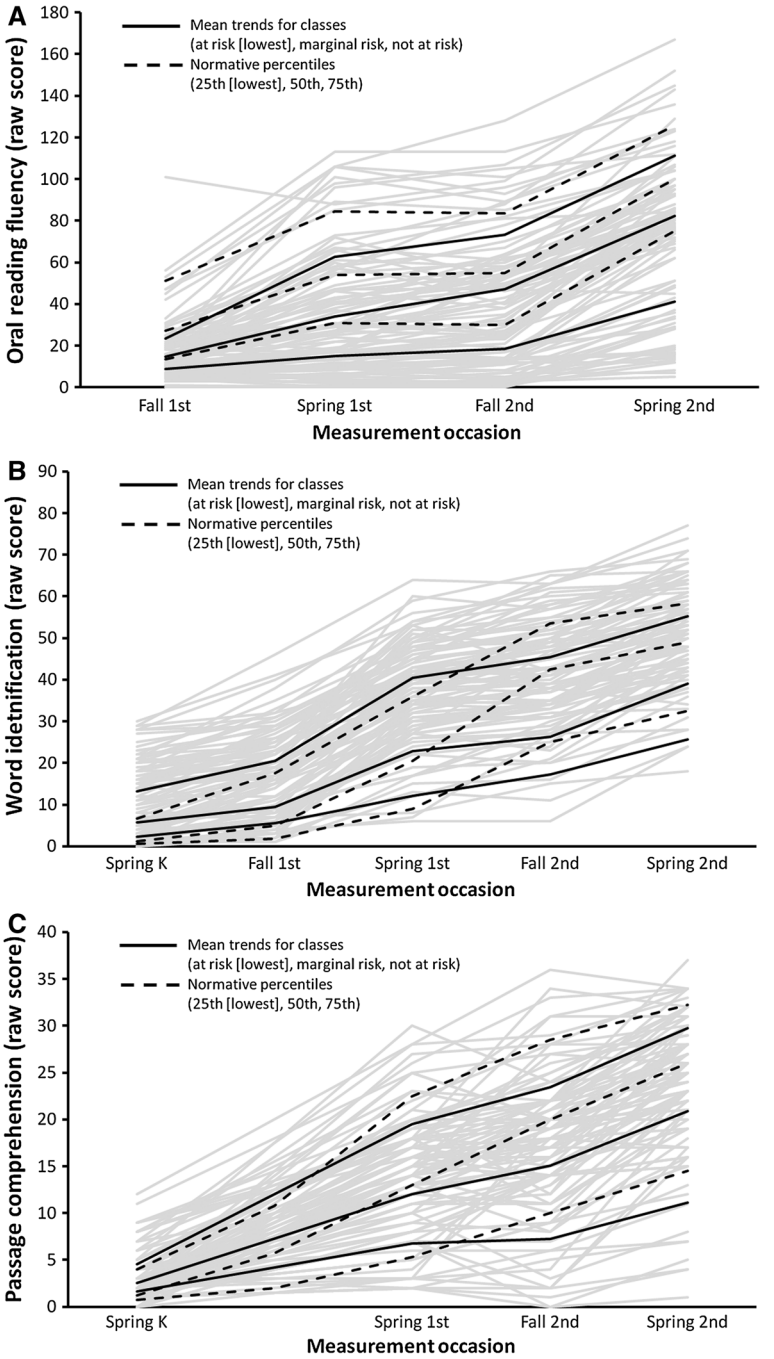
Oral reading fluency

Results for the regression model predicting oral reading fluency scores at the end of second grade are shown in Table 4. For this and subsequent models, findings are conditional in that the relations with outcome measures are based on controlling for other predictors in the model. Letter identification was a significant predictor of oral reading fluency ($B = 0.83$, $p < .05$, $sr^2 = 0.06$). Other significant effects were qualified by significant interactions, so we discuss them in the context of the interactions.

Intervention interacted significantly with ELL status ($B = -14.37$, $p < .05$, $sr^2 = 0.01$), sound matching ($B = 6.30$, $p < .05$, $sr^2 = 0.02$), and PPVT ($B = -0.63$, $p < .05$, $sr^2 = 0.03$) in predicting oral reading fluency. These significant interactions were followed up with analyses of simple slopes (Aiken & West, 1991). For ELL status, simple slopes showed that it was a significant predictor in the SDI intervention ($B = 15.90$, $p < .05$), but not in the ERI intervention ($B = 1.52$, $p > .05$). In the SDI intervention, ELL students performed better than non-ELL students on oral reading fluency; in the ERI intervention, ELL status was not a significant predictor of oral reading fluency. For sound matching, simple slopes showed that it was not a significant predictor in the SDI intervention ($B = -2.53$, $p > .05$), but it was a significant predictor in the ERI intervention ($B = 3.78$, $p < .05$; see Fig. 2). In the SDI intervention, sound matching did not predict oral reading fluency; in the ERI intervention, higher scores on sound matching at the beginning of kindergarten were associated with higher levels of oral reading fluency at the end of second grade. For PPVT, simple slopes showed that it was a significant predictor in the SDI intervention ($B = 0.53$, $p < .05$), but not in the ERI intervention ($B = -0.10$, $p > .05$; see Fig. 3). In the SDI intervention, higher PPVT scores at the beginning of kindergarten were associated with higher scores on oral reading fluency at the end of second grade; in the ERI intervention PPVT did not significantly predict oral reading fluency.

Word identification

Results for regression models predicting word identification scores at the end of second grade are shown in Table 4. Significant predictors of word identification were gender ($B = -6.96$, $p < .05$, $sr^2 = 0.01$; girls scored higher than boys), age ($B = -14.21$, $p < .05$, $sr^2 = 0.01$), and letter identification ($B = 0.79$, $p < .05$, $sr^2 = 0.05$). Other significant effects were qualified by significant interactions, so we discuss them in the context of the interactions.


A
Oral reading fluency (raw score)
180
160
140
120
100
80
60
40
20
0
Mean trends for classes
(at risk [lowest], marginal risk, not at risk)
Normative percentiles
(25th [lowest], 50th, 75th)
Fall 1st
Spring 1st
Fall 2nd
Spring 2nd
Measurement occasion
B
Word idetnification (raw score)
90
80
70
60
50
40
30
20
10
0
Mean trends for classes
(at risk [lowest], marginal risk, not at risk)
Normative percentiles
(25th [lowest], 50th, 75th)
Spring K
Fall 1st
Spring 1st
Fall 2nd
Spring 2nd
Measurement occasion
C
Passage comprehension (raw score)
40
35
30
25
20
15
10
5
0
Mean trends for classes
(at risk [lowest], marginal risk, not at risk)
Normative percentiles
(25th [lowest], 50th, 75th)
Spring K
Spring 1st
Fall 2nd
Spring 2nd
Measurement occasion

◀**Fig. 1** Raw scores across measurement occasion for oral reading fluency (**a**), word identification (**b**), and passage comprehension (**c**). *Gray lines* in each panel are for a random sample of 100 students who were present at all measurement occasions. *Solid black lines* show means for each of the three groups: the lowest is at risk, the middle is marginal risk, and the highest is not at risk. *Dashed black lines* show raw scores corresponding to normative percentile scores: the lowest is the 25th percentile, the middle is the 50th percentile, and the highest is the 75th percentile

Intervention interacted significantly with ELL status ($B = -14.72$, $p < .05$, $sr^2 = 0.01$), sound matching ($B = 5.80$, $p < .05$, $sr^2 = 0.01$), and PPVT ($B = -0.45$, $p < .05$, $sr^2 = 0.01$). These significant interactions were followed up with analyses of simple slopes. For ELL status, simple slopes showed that it was a significant predictor in the SDI intervention ($B = 15.61$, $p < .05$), but not in the ERI intervention ($B = 0.89$, $p > .05$). For sound matching, simple slopes showed that it was not a significant

Table 4 Student characteristics and baseline score predictors of oral reading fluency, word identification, and passage comprehension at end of second grade (or last available measurement)

Predictor	Oral reading fluency		Word identification		Passage comprehension	
	Coefficient	Semipartial r^2	Coefficient	Semipartial r^2	Coefficient	Semipartial r^2
Gender (male = 1)	−6.08	0.01	−6.96*	0.01	−5.95*	0.01
Age	8.23	0.00	−14.21*	0.01	−21.78*	0.03
Speech/language eligible	−0.18	0.00	−2.89	0.00	2.55	0.00
ELL	15.90*	0.03	15.61*	0.02	13.57*	0.02
Letter ID	0.83*	0.06	0.79*	0.05	0.64*	0.04
Sound matching	−2.53	0.01	−1.79	0.00	−3.37*	0.01
Rapid object naming	1.35	0.01	1.15	0.01	1.67*	0.02
PPVT	0.53*	0.03	0.53*	0.03	0.59*	0.04
Intervention (ERI = 1)	3.67	0.00	3.20	0.00	−0.10	0.00
Intervention × gender	3.05	0.00	6.85	0.00	5.92	0.00
Intervention × age	1.32	0.00	1.01	0.00	8.37	0.00
Intervention × speech/lang	−1.96	0.00	−6.82	0.00	−7.96	0.00
Intervention × ELL	−14.37*	0.01	−14.72*	0.01	−8.95	0.01
Intervention × letter ID	−0.13	0.00	0.00	0.00	−0.15	0.00
Intervention × sound match	6.30*	0.02	5.80*	0.01	8.45*	0.03
Intervention × rapid obj nam	0.64	0.00	0.19	0.00	0.15	0.00
Intervention × PPVT	−0.63*	0.03	−0.45*	0.01	−0.41*	0.01
Model	$R^2 = 0.27$		$R^2 = 0.40$		$R^2 = 0.48$	

Scores for predictor measures are standard scores; scores for outcome measures are percentile scores. All models also include amount of additional intervention and the state in which the study took place as covariates. Letter ID = WRMT-R/NU letter ID measure; Sound matching and rapid object naming are CTOPP measures; PPVT-III = Peabody Picture Vocabulary Test, version 3

* $p < .05$

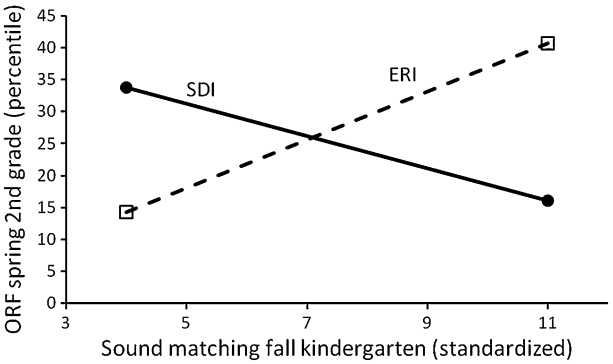


Fig. 2 Predicted oral reading fluency percentile scores in spring of second grade as a function of sound matching standard scores in fall of kindergarten and intervention

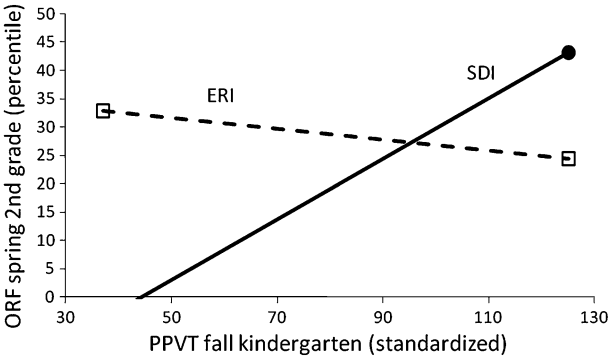


Fig. 3 Predicted oral reading fluency percentile scores in spring of second grade as a function of PPVT standard scores in fall of kindergarten and intervention

predictor in the SDI intervention ($B = -1.79, p > .05$), but it was a significant predictor in the ERI intervention ($B = 4.01$, $p < .05$). For PPVT, simple slopes showed that it was a significant predictor in the SDI intervention ($B = 0.53$, $p < .05$), but not in the ERI intervention ($B = 0.08, p > .05$). Interactions are not depicted in figures because the pattern was similar to results for oral reading fluency (see Figs. 2, 3).

Passage comprehension

Results for regression models predicting passage comprehension scores at the end of second grade are shown in Table 4. Significant predictors of passage comprehension were gender ($B = -5.95$, $p < .05$, $sr^2 = 0.01$), age ($B = -21.78$, $p < .05$, $sr^2 = 0.03$), ELL status ($B = 13.57$, $p < .05$, $sr^2 = 0.02$), letter identification ($B = 0.64$, $p < .05$, $sr^2 = 0.04$), and rapid object naming ($B = 1.67$, $p < .05$, $sr^2 = 0.02$). Other significant effects were qualified by significant interactions, so we discuss them in the context of the interactions.

Intervention interacted significantly with sound matching ($B = 8.45$, $p < .05$, $sr^2 = 0.03$) and PPVT ($B = -0.41$, $p < .05$, $sr^2 = 0.01$). These significant interactions were followed up with analyses of simple slopes. For sound matching, simple slopes showed that it was a significant negative predictor in the SDI intervention ($B = -3.37$, $p < .05$), but a significant positive predictor in the ERI intervention ($B = 5.08$, $p < .05$). For PPVT, simple slopes showed that it was a significant predictor in the SDI intervention ($B = 0.59$, $p < .05$), but not in the ERI intervention ($B = 0.18$, $p > .05$). Interactions are not depicted in figures because the pattern was similar to results for oral reading fluency (see Figs. 2, 3).

Curriculum-embedded formative measure

The performance of the curriculum-embedded formative measure as a predictor of the three outcome measures was assessed by adding it to the regression models. When added to these models, the curriculum-embedded measure was found not only to be a significant predictor of all three outcome measures but also to account for the most unique variance of any predictor in each of the three models. In predicting oral reading fluency, $B = 0.49$, $p < .05$, $sr^2 = 0.04$. In predicting word identification, $B = 0.60$, $p < .05$, $sr^2 = 0.07$. In predicting passage comprehension, $B = 0.63$, $p < .05$, $sr^2 = 0.08$. Note that the full model results are not tabled, and sr^2 values for other predictors changed from those shown in Table 4.

Discussion

The purposes of this study were to (a) examine the longitudinal reading performance of students who participated in supplemental kindergarten interventions and (b) analyze the influence and interaction of entry-level learner variables and type of intervention on end-of-second grade outcomes. We followed kindergarten students who participated in two types of supplemental reading intervention (experimental code-based and school-designed) through the end of second grade. To select predictors, we drew upon prominent variables from prior empirical research. Following we discuss findings and implications.

Performance of children who participated in supplemental kindergarten reading intervention

End-of-second-grade performance was used to classify children's risk status (at risk, marginal risk, and not at risk). Given the variability in the criteria used to identify risk across studies, we sought to standardize the risk classification process. First, children were classified as being at risk if they scored at or below the 25th percentile on the target outcome measure based on population norms. All students in the study were identified in the fall of kindergarten as being at risk of reading difficulty. End of second-grade performance indicated no statistically significant between-condition differences regarding the percentage of students by risk status. Findings of this

study indicated, however, that a combination of learner and intervention factors influenced outcomes.

Accuracy measures (word identification and passage comprehension) and the oral reading fluency measure that assessed accuracy and rate yielded markedly different risk profiles. Consistent with prior research, the majority of students who received kindergarten intervention did not evidence significant reading difficulties at the end of second grade on measures of word identification (74 %) or passage comprehension (60 %). When benchmarked against normative standards, only 10 % performed below the 25th percentile on word identification and 15 % on passage comprehension. The proportion of students identified in the present study as at risk is consistent with the results of Vellutino et al. (2006), in which 16 % of their at-risk sample in kindergarten remained at risk at the end of third grade based on the 25th percentile cut score.

In contrast, findings indicated that 55 % of students remained at risk (i.e., performed below the 25th percentile) on the oral reading fluency outcome measure. We included oral reading fluency based on its prior use in response to intervention research (Al Otaiba & Fuchs, 2006). Although prior research suggests that oral reading fluency that assesses accuracy and automaticity is a reliable predictor of reading comprehension, rigorous fluency benchmarks can result in high rates of over identification (e.g., Johnson et al., 2009) and risk rates that are much greater than nonfluency based reading measures (Denton et al., 2010). In contrast, researchers have questioned the utility of WRMT-R subtests in evaluating intervention effects due to inflated results (Fuchs, Fuchs, & Compton, 2004). As schools embrace a preventive approach to reading instruction and intervention, the discrepancy between accuracy and fluency-based measures highlights the need to improve the methods used to measure risk and response to intervention.

Building on prior empirical research, this study's findings validate the insufficiency of early intervention research for the full range of students identified as at risk of reading difficulty (e.g., Torgesen, 2000; Vellutino, Scanlon, Small, & Fanuele 2006). That is, not only did a percentage of children remain at risk of reading difficulty, their relative performance at the end of kindergarten failed to improve over time. Across measures, percentile rankings of students in the at-risk group ranged between the 14th and 31st percentile at the end of kindergarten or the beginning of first grade, and declined to the 9th through 16th percentile at the end of second grade. The importance of this finding is that, based on data at the end of kindergarten, we could have identified children whose performance was unlikely to improve without more intensive intervention.

Learner and intervention predictors of second-grade performance

Findings regarding predictors of reading achievement substantiate the complexity of defining early identification models (Barth et al., 2008). Consistent with prior predictor (e.g., Schatschneider et al., 2004) and intervention research (e.g., Al Otaiba & Fuchs, 2002; Vellutino et al., 2006), we found that letter identification assessed prior to kindergarten intervention predicted reading achievement at the end of second grade for both interventions. Specifically, the ability to recognize letter names predicted performance on all three outcome measures: word

identification, oral reading fluency, and passage comprehension. Whereas letter identification has emerged as a reliable predictor of primary-grade reading outcomes, few studies have collected these measures prior to intervention. Foulin's (2005) review of studies sheds important perspective on the role of letter naming knowledge to reading acquisition and children's response to intervention. Specifically, she concluded that letter name knowledge is foundational to reading and spelling in that it "boosts the development of major literacy-related skills, notably letter-sound knowledge and phonemic sensitivity skills, which in turn underlie the acquisition of the alphabetic principle and reading and spelling skills" (p. 17). With respect to predictors or reading achievement, findings indicate that letter name knowledge is a valid predictor of students who will require more intensive instruction.

Consistent with previous research syntheses (Nelson et al., 2003), student demographics, and specifically gender and age were significant predictors of word identification and passage comprehension. At the beginning of kindergarten, being male and older predicted lower reading outcomes at the end of second grade.

An aim of this research was to examine whether the influence of predictors varied by type of intervention. Specifically we found that the influence of three predictors (receptive vocabulary, sound matching, and English language learner) was moderated by type of intervention. Our findings corroborate those of Connor et al. (2004, 2009) regarding the interaction of child characteristics and type of intervention. Findings indicated a nuanced set of interactions that warrant further investigation. We consider these findings important as they add information that may lead to more customized interventions.

Findings from our initial predictor models indicated that receptive vocabulary was significantly associated with all second-grade outcomes, a finding that is consistent with empirical research by Vadasy et al. (2008) and Al Otaiba and Fuchs (2006). Had we not examined interactions, we would have concluded a comparable influence of vocabulary between interventions. Analysis of interaction plots indicated that for SDI, vocabulary predicted outcomes in the direction one would hypothesize. That is, students with higher entry-level vocabulary had higher end-of-second grade achievement and students with lower entry-level vocabulary had lower end-of-second grade achievement. In ERI, however, low entry-level vocabulary did not suppress end-of-second grade outcomes. In a previous study examining child-by-instruction interaction, Connor et al. (2004) found that first-graders with low initial vocabulary scores achieved greater decoding growth with more time spent in teacher-managed explicit decoding instruction early in the academic year. Our current finding suggests that children with initial low vocabulary were advantaged by instruction in the ERI that was explicit and consistent. That is, when the language of instruction is carefully engineered as in the ERI program, it may lessen the impact of entry-level vocabulary on reading outcomes.

Results of the significant ELL-by-intervention analyses indicated that on oral reading fluency and word identification outcomes, being an ELL versus a non-ELL was a significant positive predictor in SDI, but it was nonsignificant in ERI. This finding suggests that on measures that involved word reading accuracy and rate, SDI instruction benefitted ELL students more than non-ELL students while ERI was

comparably effective for both ELL and non-ELL students on the same measures. This interaction was not found on the passage comprehension measure that ostensibly placed greater requirements on language comprehension. Prior research has documented that with quality instruction ELLs can reach levels of word reading accuracy commensurate with their monolingual English-speaking peers (Lesaux, Lipka, & Siegal, 2006), yet reading comprehension levels continue to lag for ELLs, most likely due to language comprehension. Findings from the current study suggest that there is benefit of small-group reading intervention for ELL students on measures of word reading and rate, a finding that has been documented in previous studies and research syntheses (Gersten et al., 2007; O'Connor, Bocian, Beebe-Frankenberger, & Linklater, 2010). A logical explanation for the differential effect of SDI among ELLs and their English-speaking peers is that the early code-based difficulties of ELLs may be more malleable than the code-based difficulties of their peers whose difficulties may be rooted in phonological and alphabetic processing difficulties. SDI was effective for ELLs who needed early code-based instruction but not for English-speaking students who needed more systematic intervention. Though the sources of ELLs reading difficulties are far from clear, research (e.g., Kieffer & Vukovic, 2012) suggests that while code-based instruction is necessary, it is not sufficient and linguistic comprehension may play a larger role in explaining variation in ELL performance than code-based components. While both SDI and ERI code-based instruction were beneficial for ELLs, reading interventions that impact reading comprehension will need to address the role of language more explicitly.

Interaction analyses further revealed that the influence of the sound-matching phonological awareness measure interacted with intervention. Specifically, the direction of relations differed by intervention. Higher sound-matching scores in ERI predicted higher end-of-second performance on all outcomes. Conversely, higher entry-level scores in SDI predicted poorer end of second-grade performance in SDI. The finding that phonological awareness predicted response in a code-based intervention (ERI) is consistent with previous intervention research (Al Otaiba & Fuchs, 2006; Vellutino et al., 2006). The negative relation between phonological awareness and reading outcomes in SDI was surprising and warrants further study. Schatschneider et al. (2004) noted that assessments of phonological awareness collected at the beginning of kindergarten are less reliable than those collected in the middle and end of kindergarten. Moreover, they concluded that the relative importance and predictive validity of phonological awareness may depend on the types of measures used. Our reliance on an early and less complex measure may have contributed to the variability of findings.

A secondary analysis involving only children in the experimental intervention showed that performance on the curriculum-embedded test, a formative measure, was uniquely predictive of all second-grade outcome measures. This test assesses a child's mastery of content taught in the first 8 weeks of intervention. A previous study (Oslund et al., 2012) found that this measure was a predictor of kindergarten outcomes, and the current finding extends its practical utility. Moreover, this finding provides supporting evidence for the RTI practice guide panel's (Gersten et al., 2009) recommendation to monitor the progress of students in supplemental

intervention. As an early indicator of response to intervention, curriculum-embedded measures may provide important information regarding the focus and content of more intensive intervention.

Study limitations

We want to acknowledge several limitations of this study. First, and perhaps foremost, longitudinal followup studies are inherently complex and often incomplete. Attrition in our sample was 31 % and while between-group attrition did not differ, students who were not present at posttests could conceivably have influenced outcomes. In addition, to understand factors associated with student achievement in the long term, one would ideally document the quantity and quality of instruction that transpired between the end of intervention and the outcome measurement period. We know that schools provide a continuum of intervention. To fully explain student achievement, we would need measures of classroom instruction and the quality of that instruction as well as amounts and type of intervention. Given the number of students who were English language learners, information on the type of instruction received would likewise be useful. There is a clear need for intensive, descriptive, longitudinal research that accurately represents and records the full range and quality of instruction and interventions that transpire over time.

A methodological limitation relates to the predictor set. Specifically, we did not include the full complement of behavioral, linguistic, and cognitive processing variables used in previous studies. We did, however, include predictors derived from prominent theories of reading difficulties. Inconsistencies in the predictive validity of phonological awareness may be due to our reliance on one of the least complex measures of phonological awareness. Using only sound matching limits the generalizability of findings and may explain why phonological awareness was not sufficiently robust across intervention groups. We elected to use sound matching as there were norms for students in our sample; however, augmenting this measure with one with greater discriminatory power, such as sound blending, may extend our understanding of the entry-level predictive validity of phonemic awareness.

A final factor that warrants discussion is our method of quantifying response. Some would argue that the 25th percentile is too low a threshold of reading competence necessary to meet the demands of complex text. Others that the 25th percentile identifies more children than schools can adequately address. While the field is not clear on what type of criteria should be used to establish student response (Barth et al., 2008; Cirino et al., 2009), we consider it important to increase the confidence of our decisions regarding a student's risk status. Were the methods we used to classify students as at risk too lenient or the confidence interval to classify children as not at risk too stringent?

Without doubt, risk and response can be determined along a number of dimensions. We feel confident that students who were classified as being not at risk performed above the 25th percentile. Present and prior research has documented the moderate to strong correlations between word identification, oral reading fluency, and passage comprehension. However, the criteria for our performance groups identified more than half of

the students as being at risk at the end of second grade, a figure that would likely exceed schools' resources and ability to supplemental intervention.

Implications for research and practice

Evidence from this study extends our understanding of the later outcomes of students with early reading risk who receive intervention in kindergarten (Al Otaiba & Fuchs, 2006; Barth et al., 2008; O'Connor, Fulmer, Harty, & Bell, 2005; Simmons et al., 2008; Vellutino et al., 2006). Many students in both experimental and school-designed interventions evidenced average to above-average performance on important transfer measures 2 years post intervention (Compton et al., 2005). Specifically, approximately 85 % of students performed solidly above average (60th percentile) on word identification and passage comprehension. Based on findings, however, neither kindergarten intervention nor typical instruction provided in first or second grade sufficiently prepared the majority of students to read grade-level text fluently. Performance trends revealed that children in the at-risk group made minimal fluency progress with end of first- and second-grade oral reading fluency performance averaging around the 10th percentile.

While significant strides have been made in identifying children who are likely to experience difficulty, findings from this study suggest the need for further information. Findings from predictor analyses corroborate the validity of letter identification as a predictor and the complexity of identifying and interpreting the differential influence of predictors by intervention. Study findings suggest that the influence of particular variables vary by type of intervention and further research is needed that examines how different types of instruction can moderate the influence of entry-level skills and characteristics. For the experimental intervention group in this study, the curriculum-embedded measure was a robust predictor for all outcome measures, further validating the utility of measures that assess students' early response to intervention through mastery of specific content.

Finally, this study reinforces the need for future research on interventions for children who fail to respond during kindergarten and demonstrate prolonged reading difficulty. Approximately 15 % of students performed below the 25th percentile on the most liberal measures administered. To more effectively design and deliver instruction for these students, further research may build on the factors that reliably predict outcomes across intervention (e.g., letter identification) and continue to examine the interaction of learner characteristics and intervention attributes.

References

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park, CA: Sage.

Al Otaiba, S., Folsom, J. S., Schatschneider, C., Wanzek, J., Greulich, L., Meadows, J., et al. (2011). Predicting first-grade reading performance from kindergarten response to tier 1 instruction. *Exceptional Children, 77*, 453–470.

Al Otaiba, S., & Fuchs, D. (2002). Characteristics of children who are unresponsive to early literacy intervention: A review of the literature. *Remedial and Special Education, 23*, 300–316. doi: 10.1177/07419325020230050501.

Al Otaiba, S., & Fuchs, D. (2006). Who are the young children for whom best practices in reading are ineffective? An experimental and longitudinal study. *Journal of Learning Disabilities, 39*, 414–431. doi:10.1177/00222194060390050401.

Barth, A. E., Stuebing, K. K., Anthony, J. L., Denton, C. A., Mathes, P. G., Fletcher, J. M., et al. (2008). Agreement among response to intervention criteria for identifying responder status. *Learning and Individual Differences, 18*, 296–307. doi:10.1016/j.lindif.2008.04.004.

Carrow-Woolfolk, E. (1995). *Oral and Written Language Scales: Listening comprehension and oral expression*. Cedar Pines, MN: American Guidance Service.

Cavanaugh, C. L., Kim, A.-H., Wanzek, J., & Vaughn, S. (2004). Kindergarten reading interventions for at-risk students: Twenty years of research. *Learning Disabilities: A Contemporary Journal, 2*(1), 9–21.

Chard, D. J., Stoolmiller, M., Harn, B. A., Wanzek, J., Vaughn, S., Linan-Thompson, S., et al. (2008). Predicting reading success in a multilevel schoolwide reading model: A retrospective analysis. *Journal of Learning Disabilities, 41*, 174–188. doi:10.1177/0022219407313588.

Cirino, P. T., Vaughn, S., Linan-Thompson, S., Cardenas-Hagen, E., Fletcher, J. M., & Francis, D. J. (2009). One-year follow-up outcomes of Spanish and English interventions for English language learners at risk for reading problems. *American Educational Research Journal, 46*, 744–781. doi:10.3102/0002831208330214.

Compton, D. L., Olinghouse, N. G., Elleman, A., Vining, J., Appleton, A. C., Vail, J., et al. (2005). Putting transfer back on trial: Modeling individual differences in the transfer of decoding-skill gains to other aspects of reading acquisition. *Journal of Educational Psychology, 97*, 55–69. doi:10.1037/0022-0663.97.1.55.

Connor, C. M., Morrison, F. J., & Katch, E. L. (2004). Beyond the reading wars: The effect of classroom instruction by child interactions on early reading. *Scientific Studies of Reading, 8*, 305–336.

Connor, C. M., Piasta, S. B., Fishman, B., Glasney, S., Schatschneider, C., Crowe, E., et al. (2009). Individualizing student instruction precisely: Effects of child × instruction interactions on first graders' literacy development. *Child Development, 80*, 77–100.

Coyne, M. D., Kame'enui, E. J., Simmons, D. C., & Harn, B. A. (2004). Beginning reading intervention as inoculation or insulin: First-grade reading performance of strong responders to kindergarten intervention. *Journal of Learning Disabilities, 37*, 90–104. doi:10.1177/00222194040370020101.

Coyne, M. D., Little, M., Rawlinson, D., Simmons, D., Kwok, O., Kim, M., Simmons, L., Hagan-Burke, S., & Civetelli, C. (2013). Replicating the impact of a supplemental beginning reading intervention: The role of instructional context. *Journal of Research on Educational Effectiveness, 6*, 123. doi:10.1080/19345747.2012.706694.

Denton, C. A., Kethley, C., Nimon, K., Kurz, T. B., Mathes, P. G., Shih, M., et al. (2010). Effectiveness of a supplemental early reading intervention scaled up in multiple schools. *Exceptional Children, 76*, 394–416.

Dunn, L. M., & Dunn, L. M. (1997). *Peabody picture vocabulary test—Third edition (PPVT-III)*. Bloomington, MN: Pearson Assessments.

Foorman, B. R., Breier, J. I., & Fletcher, J. M. (2003). Interventions aimed at improving reading success: An evidence-based approach. *Developmental Neuropsychology, 24*, 613–639.

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.

Foulin, J. N. (2005). Why is letter-name knowledge such a good predictor of learning to read? *Reading and Writing, 18*, 129–155. doi:10.1007/s11145-004-5892-2.

Fuchs, D., Fuchs, L. S., & Compton, D. L. (2004). Identifying reading disabilities by responsiveness-to-instruction: Specifying measures and criteria. *Learning Disability Quarterly, 27*, 216–228. doi:10.2307/1593674.

Gersten, R., Baker, S. K., Shanahan, T., Linan-Thompson, S., Collins, P., & Scarcella, R. (2007). Effective literacy and English language instruction for English learners in the elementary grades: A practice guide (NCEE 2007-4011). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from http://ies.ed.gov/ncee/wwc/publications/practiceguides.

Gersten, R., Compton, D., Connor, C. M., Dimino, J., Santoro, L., Linan-Thompson, S., et al. (2009). Assisting students struggling with reading: Response to Intervention and multi-tier intervention in the primary grades. A practice guide (NCEE 2009-4045). Washington, DC: U.S. Department of Education, National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences. Retrieved from http://ies.ed.gov/ncee/wwc/publications/practiceguides/.

Good, R. H., I. I. I., & Kaminski, R. A. (Eds.). (2002). *Dynamic indicators of basic early literacy skills* (6th ed.). Eugene, OR: Institute for the Development of Educational Achievement.

Hogan, T. P., & Thomson, J. M. (2010). Epilogue to *Journal of Learning Disabilities* special edition, "Advances in the early detection of reading risk": Future advances in the early detection of reading risk: Subgroups, dynamic relations, and advanced methods. *Journal of Learning Disabilities, 43*, 383–386. doi:10.1177/0022219410369077.

Johnson, E. S., Jenkins, J. R., Petscher, Y., & Catts, H. W. (2009). How can we improve the accuracy of screening instruments? *Learning Disabilities Research and Practice, 24*, 174–185. doi:10.1111/j.1540-5826.2009.00291.x.

Kieffer, M. J., & Vukovic, R. K. (2012). Components and context: Exploring sources of reading difficulties for language minority learners and native English speakers in urban schools. *Journal of Learning Disabilities, 45*, 433–452. doi:10.1177/0022219411432683.

Lesaux, N. K., Lipka, O., & Siegel, L. S. (2006). Investigating cognitive and linguistic abilities that influence the reading comprehension skills of children from diverse linguistic backgrounds. *Reading and Writing: An Interdisciplinary Journal, 19*, 99–131. doi:10.1007/s11145-005-4713-6.

Nelson, J. R., Benner, G. J., & Gonzalez, J. (2003). Learner characteristics that influence the treatment effectiveness of early literacy interventions: A meta-analytic review. *Learning Disabilities Research and Practice, 18*, 255–267. doi:10.1111/1540-5826.00080.

O'Connor, R. E., Bocian, K., Beebe-Frankenberger, M., & Linklater, D. L. (2010). Responsiveness of students with language difficulties to early intervention in reading. *The Journal of Special Education, 43*, 220–235. doi:10.1177/0022466908317789.

O'Connor, R. E., Fulmer, D., Harty, K. R., & Bell, K. M. (2005). Layers of reading intervention in kindergarten through third grade: Changes in teaching and student outcomes. *Journal of Learning Disabilities, 38*, 440–455. doi:10.1177/00222194050380050701.

Oslund, E. L., Hagan-Burke, S., Taylor, A. B., Simmons, D. C., Simmons, L., Kwok, O., Johnson, C., & Coyne, M. D. (2012). Predicting kindergarteners' response to early reading intervention: An examination of progress-monitoring measures. *Reading Psychology, 33*, 78–103. doi:10.1080/02702711.2012.630611.

Pearson/Scott Foresman. (2004). *Scott Foresman sidewalks: Early reading intervention*. Glenview, IL: Pearson Education, Inc.

Scammaca, N., Vaughn, S., Roberts, G., Wanzek, J., & Torgesen, J. K. (2007). *Extensive reading interventions in grades K-3: From research to practice*. Portsmouth, NH: RMC Research Corporation, Center on Instruction.

Schatschneider, C., Fletcher, J. M., Francis, D. J., Carlson, C. D., & Foorman, B. R. (2004). Kindergarten prediction of reading skills: A longitudinal comparative analysis. *Journal of Educational Psychology, 96*, 265–282. doi:10.1037/0022-0663.96.2.265.

Simmons, D. C., Coyne, M. D., Kwok, O., McDonagh, S., Harn, B. A., & Kame'enui, E. J. (2008). Indexing response to intervention: A longitudinal study of reading risk from kindergarten through third grade. *Journal of Learning Disabilities, 41*, 158-173. doi:10.1177/0022219407313587

Simmons, D. C., Coyne, M. D., Hagan-Burke, S., Kwok, O., Simmons, L. E., Johnson, C., Zou, Y., Taylor, A. B., McAlenney, A. L., Ruby, M., & Crevecoeur, Y. C. (2011). Effects of supplemental reading interventions in authentic contexts: A comparison of kindergarteners response. *Exceptional Children, 77*, 207228.

Torgesen, J. K. (2000). Individual differences in response to early interventions in reading. *Learning Disabilities Research and Practice, 15*, 55–64.

Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (1999). *Test of word reading efficiency*. Austin, TX: Pro-Ed.

University of Oregon Center on Teaching and Learning. (2002). *Assessment committee: Analysis of reading assessment measures.* Retrieved from https://dibels.uoregon.edu/techreports/dibels_5th_ed.pdf.

Vadasy, P. F., Sanders, E. A., & Abbott, R. D. (2008). Effects of supplemental early reading intervention at 2-year follow up: Reading skill growth patterns and predictors. *Scientific Studies of Reading, 12*, 51–89. doi:10.1080/10888430701746906.

Vadasy, P. F., Sanders, E. A., & Peyton, J. A. (2006). Code-oriented instruction for kindergarten students at risk for reading difficulties: A randomized field trial with paraeducator implementers. *Journal of Educational Psychology, 98*, 508–528. doi:10.1037/0022-0663.98.3.508.

Vellutino, F. R., Scanlon, D. M., Sipay, E. R., Small, S. G., Pratt, A., Chen, R. S., et al. (1996). Cognitive profiles of difficult-to-remediate and readily remediated poor readers: Early intervention as a vehicle

for distinguishing between cognitive and experiential deficits as basic causes of specific reading disability. *Journal of Educational Psychology, 88*, 601–638. doi:10.1037//0022-0663.88.4.601.

Vellutino, F. R., Scanlon, D. M., Small, S., & Fanuele, D. P. (2006). Response to intervention as a vehicle for distinguishing between children with and without reading disabilities: Evidence for the role of kindergarten and first-grade interventions. *Journal of Learning Disabilities, 39*, 157–169. doi: 10.1177/00222194060390020401.

Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1999). *Comprehensive test of phonological processing*. Austin, TX: Pro-Ed.

Woodcock, R. W. (1987, 1998). *Woodcock reading mastery tests—Revised/Normative Update*. Bloomington, MN: Pearson Assessments.