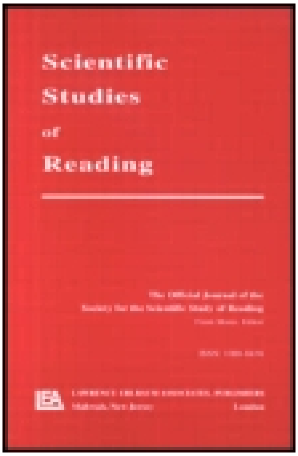
Sound Partners

Efficacy of Supplemental Phonics-Based Instruction for Low-Skilled First Graders: How Language Minority Status and Pretest Characteristics Moderate Treatment Response

Patricia F. Vadasy and Elizabeth A. Sanders

VOYAGER SOPRIS LEARNING®



Scientific Studies of Reading
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/hssr20

Efficacy of Supplemental Phonics-Based Instruction for Low-Skilled First Graders: How Language Minority Status and Pretest Characteristics Moderate Treatment Response

Patricia F. Vadasy [a] & Elizabeth A. Sanders [a]
[a] Washington Research Institute, Seattle, USA
Published online: 24 Feb 2011.

SCIENTIFIC STUDIES OF READING, *15*(6), 471–497

ISSN: 1088-8438 print / 1532-799X online
DOI: 10.1080/10888438.2010.501091

Efficacy of Supplemental Phonics-Based Instruction for Low-Skilled First Graders: How Language Minority Status and Pretest Characteristics Moderate Treatment Response

Patricia F. Vadasy and Elizabeth A. Sanders
Washington Research Institute, Seattle

We examined the efficacy of 20 weeks of individual supplemental phonics-based instruction for language minority (LM) and non-LM first graders. Students were designated LM if the primary home language was not English (otherwise non-LM). Those performing in the bottom half of their classroom LM/non-LM group in letter knowledge and phonological awareness were randomly assigned to treatment and control conditions. Treatment included alphabetics, decoding, and oral reading practice. Results showed that treatment students ($n = 93$) outperformed controls ($n = 94$) on 5 of the 6 posttests; however, LM students exhibited lower treatment response on passage reading fluency. Pretest word reading did not moderate treatment response, and LM students with greater baseline vocabulary showed greater treatment response on posttest word reading and spelling.

A robust and well-differentiated body of research supports the effectiveness of classroom-based early reading interventions (e.g., Blachman, Ball, Black, & Tangel, 1994; Foorman, Francis, Fletcher, Schatschneider, & Mehta, 1998; Hatcher, Hulme, & Ellis, 1994), as well as non-classroom-based supplemental interventions implemented in small groups (e.g., Mathes et al., 2005; Torgesen, Wagner, Rashotte, Herron, & Lindamood, 2009; Watson & Hempenstall, 2008) or one-to-one (e.g., O'Connor, 2000; Torgesen, Wagner, & Rashotte, 1997; Vellutino

Correspondence should be sent to Patricia F. Vadasy, Washington Research Institute, 150 Nickerson Street, Suite 305, Seattle, WA 98109. E-mail: pvadasy@wri-edu.org

et al., 1996). Both small-group and individual supplemental interventions have included teacher-tutors (e.g., Center, Wheldall, Freeman, Outhred, & McNaught, 1995; Torgesen et al., 1997), paraeducators (e.g., Brown, Morris, & Fields, 2005; Ehri, Dreyer, Flugman, & Gross, 2007; Vadasy, Sanders, & Abbott, 2008), and other types of tutors (e.g., Allor & McCathren, 2004; Baker, Gersten, & Keating, 2000; Gelzheiser, 2005).

Although there is widespread support for the importance of early reading intervention, less is known about the effects of early reading interventions for children with limited English language proficiency. Most early reading interventions focus on the contribution of decoding skill rather than listening comprehension to reading comprehension (see the "simple view of reading"; Gough & Tunmer, 1986). Early reading interventions for native English speaking children that have a strong phonics emphasis (such as that tested in the current study) are associated with moderate to high effect sizes in word-level reading skills and with somewhat smaller effect sizes in text-level fluency and comprehension skills (Ehri, Nunes, Stahl, & Willows, 2001; Slavin, Lake, Davis, & Madden, 2009). Considerable research with native English speakers on early risk factors for reading problems has shown that first-grade phonological awareness, alphabetic knowledge, and word reading are each strong first-grade predictors of later reading abilities (e.g., Jenkins, Hudson, & Johnson, 2007). In addition, studies of native English speakers have demonstrated that oral language skills, including vocabulary knowledge (often used as a proxy for general oral language skills), play a larger role in reading comprehension in later grades (e.g., O'Connor & Jenkins, 1999; Vellutino, Tunmer, Jaccard, & Chen, 2007; Whitehurst & Lonigan, 1998).

How to identify language minority (LM) students who are likely to have difficulty learning to read and could benefit from early reading intervention is less clear. Disaggregated norms for LM students are often not available for widely used screening measures, and it is difficult for educators to evaluate the role of English language proficiency in LM children's reading development and response to reading intervention. In several studies that have tested the generalizability of the simple view for LM students (Gough & Tunmer, 1986; Hoover & Gough, 1990), both English oral language and English word reading skills were shown to uniquely contribute to English reading comprehension. Gottardo and Mueller (2009), in turn, found that first-grade English phonological awareness and oral language (receptive vocabulary and oral cloze) each uniquely predicted second-grade English reading comprehension. Further, Proctor, August, Carlo, and Snow (2006) reported that English listening comprehension and vocabulary knowledge were more predictive of reading comprehension than decoding skills for LM fourth graders. We investigate whether first-grade LM students' pretest English vocabulary is associated with response to a supplemental early reading intervention.

READING INTERVENTIONS FOR LM STUDENTS

LM students often enter school with lower levels of early phonological, word reading, and oral language skills compared with their native English-speaking peers (August & Shanahan, 2006; Betts, Bolt, Decker, Muyskens, & Marston, 2009; Lesaux, 2006). A growing body of studies shows that when young LM students are provided with strong classroom or supplemental instruction in early reading skills, LM students “catch up” in these early reading skills (e.g., Chiappe & Siegel, 2006; Droop & Verhoeven, 2003; Lipka & Siegel, 2007). Similarly, research has shown that LM students who receive research-based classroom reading intervention attain word reading skills similar to non-LM students but may lag behind in syntactic knowledge and comprehension (Lesaux & Siegel, 2003; Verhoeven, 1990). Finally, two expert reviews suggest that interventions with a strong phonics emphasis will benefit LM students. The first was a qualitative review by Shanahan and Beck (2006) of five studies (those that met review criteria for the National Literacy Panel on Language-Minority Children and Youth) that investigated phonemic awareness and phonics instruction for LM students. The results indicated that early phonics-based interventions are beneficial for LM students’ word-level skills. The second review, conducted by the U.S. Department of Education, Institute of Education Sciences (2007), examined three elementary reading interventions for LM students that employed differing delivery models. Two of these studies (Denton, Anthony, Parker, & Hasbrouck, 2004; Gunn, Biglan, Smolkowski, & Ary, 2000) demonstrated that supplemental systematic word-level instruction produced positive outcomes in word reading, although effect sizes for word reading and comprehension were smaller than seen in similar interventions for non-LM students.

To our knowledge, only one prior study has directly compared LM and non-LM students’ response to reading intervention (others have tested vocabulary interventions). Specifically, O’Connor, Bocian, Beebe-Frankenberger, and Linklater (2010) tested a half-year, supplemental, small-group reading intervention with 78 LM and non-LM kindergarteners who were below the 16th percentile (1 *SD* below average) in English vocabulary and who performed at or below the “some-risk” levels on two alphabetic fluency tasks. Children were randomly assigned to treatment or delayed treatment (control) conditions within schools. Results showed treatment effects on all posttests (alphabetic and phonological tasks), no significant effects of LM status, and no Treatment × LM interactions. Further, pretest vocabulary, but not language proficiency, was significantly correlated with all posttests. Finally, their results showed that pretest vocabulary (although restricted) did not predict treatment responsiveness. In the current study, we test whether English vocabulary predicts LM student outcomes as well as whether vocabulary moderates LM students’ treatment response.

CURRENT STUDY

In the present study, we test the treatment effects of a one-to-one, supplemental, phonics-based early reading intervention with both LM and non-LM first graders. This intervention has been used in prior research with primarily non-LM first-grade students who performed in the bottom quartile of word reading in the fall of first grade (Jenkins, Peyton, Sanders, & Vadasy, 2004; Vadasy, Jenkins, Antil, Wayne, & O'Connor, 1997a, 1997b) and aligns with research on effective instructional practices to develop early literacy skills in low-performing primary-age students (e.g., Blachman et al., 1994; Hatcher et al., 2006; Schneider, Roth, & Ennemoser, 2000). Our interest in testing whether supplemental reading intervention benefits lower skilled LM students stems from the need to improve reading skills for the growing number of LM students in the United States who are at elevated risk for reading difficulties (August & Hakuta, 1998; Lesaux & Geva, 2006). Between 1979 and 2007 the number of LM students in the United States increased from 9% to 20% of the school-age population (Planty et al., 2009). Enrollments in the proportion of school-age students who are immigrants with a home language other than the dominant language have likewise increased in many European countries and Canada (Canadian Education Statistics Council, 2007; Eurydice, 2004). These LM and immigrant students are more likely to be poor (Heath, Rothon, & Kilpi, 2008; University of California Linguistic Minority Research Institute, 2007), and both LM and low-income status are associated with increased risk of reading difficulties (Snow, Burns, & Griffin, 1998). Further, both LM and low-income students are more likely to attend low-achieving schools (Capps et al., 2005; Fry, 2008), and early social inequalities in academic outcomes increase as students advance through the elementary grades (Entwisle, Alexander, & Olson, 1997; Kieffer, 2008; Phillips, Crouse, & Ralph, 1998).

RESEARCH QUESTIONS

The current study adds to the literature on early reading intervention for LM and non-LM students by testing the efficacy of phonics-based supplemental reading intervention with first graders performing in the lower half of their classroom LM/non-LM group. Specifically, we screened all first-grade LM and non-LM students in each classroom on alphabetic and phonological awareness measures and selected the bottom-performing half of each group for study participation. Students may thus be considered at *relative risk* for reading difficulties because they were selected from within LM/non-LM groups, within classrooms. Eligible students were then randomly assigned to treatment (supplemental reading intervention) or control (no intervention; classroom instruction only) conditions. This study design assured adequate representation of both LM and non-LM children in treatment and control groups within a single sample, enabling us to have a balanced test of the interaction between LM and treatment status on student outcomes. Further, we employ multilevel modeling to account for school and classroom variance in all of our analyses.

Of importance, the current study explicitly tests treatment moderators—particularly the joint effect of pretest word reading and treatment on posttests. Finally, as an auxiliary research question, we test the interaction between treatment and pretest English receptive vocabulary on posttests for LM students (we had too few non-LM students with equivalent lower vocabulary levels to allow us to include non-LM students in these analyses). Results from this research question can help inform identification of LM students who may benefit most from early literacy intervention. Our research questions are specifically as follows.

Primary RQ1: What are the effects of first-grade supplemental reading intervention (treatment) and LM status on student outcomes, and do treatment effects depend on LM status?

Primary RQ2: Does pretest word reading moderate treatment response, and if so, does this relationship depend on LM status?

Auxiliary RQ: For LM students in particular, what are the moderating effects of pretest receptive (English) vocabulary on student outcomes?

METHOD

Participants

Initial sample. In September of 2007–08, we invited all students in first-grade classrooms at 13 urban public elementary schools, known for relatively large proportions of language minority (LM) enrollment and large numbers of students performing below state proficiency levels in reading, to participate in our research study. Students were classified as LM if the student's parent reported *the primary home language as other than English on the student's school registration record*. Students receiving extra services, including special education and English Learner services, were not excluded from participation.

Of the 903 students invited via letters sent home to parents, 553 students had consents returned (282 were LM). All invitations and consent forms were sent home in English, and for the top 10 most frequent languages in the district we also sent translated invitations and consent forms. In our initial sample, there were at least 28 languages represented (some were other African languages not specified); the top 5 most frequent languages, in rank order, were Spanish (63% of LM students), Vietnamese (9%), Chinese (6%), Somali (5%), and Tagalog (2%).

Due to insufficient student sample sizes for random assignment to experimental conditions (within LM/non-LM groups within classrooms), 15 classrooms (and thus two schools) were removed from study participation prior to screening. In addition, some students had moved from their schools prior to, or were persistently absent during, screening. As such, 399 students (214, or 54%, were LM) were screened. Screening occurred in September/October, and assessments included measures of alphabetic knowledge (number of letter sounds and letter names produced out of 52 randomly ordered uppercase English letters; Fuchs et al., 2001) and phonological awareness (Sound Matching subtest from the Comprehensive Test of Phonological Processing (Wagner, Torgesen, & Rashotte, 1999). Figure 1 illustrates the flow of participants from recruitment to posttest.

Sampling design and treatment assignment. We randomly assigned children to treatment and control groups within classrooms, within LM/non-LM group. To accomplish this we first separated students by classroom; then, we separated LM and non-LM students within each classroom. Next, we computed a composite z score for each student, based on the mean z score of the three screening measures, within LM/non-LM group, within classroom. Students were then rank ordered, and students in the lower half of their classroom's LM/non-LM group were randomly assigned to treatment (supplemental phonics-based tutoring) or control (regular classroom instruction, no tutoring) conditions. We employed this relative-risk sampling strategy primarily to assure adequate and roughly equal group sizes for testing whether treatment effects were moderated by LM status. This sampling method resulted in the inclusion of students who would not typically be identified for early intervention. We examine the issue of risk status by explicitly testing whether pretest word reading performance (a risk factor used in similar interventions; e.g., Jenkins et al., 2007; Lesaux & Siegel, 2003) is associated with treatment response.

Final sample. Fifteen students moved (7 treatment and 8 controls) and 6 treatment students (5 LM and 1 non-LM) were randomly selected out of the study (within classroom) due to insufficient tutoring resources at their school. After attrition, the final sample included 93 treatment students (48 LM students) and 94 controls (50 LM students) from 29 classrooms across 11 schools. These schools had student enrollments averaging 84% minority, 74% free or reduced lunch, 35% transitional bilingual, and 16% special education during the intervention year. All of the participating schools were designated Title I. Table 1 summarizes student characteristics for each condition by LM/non-LM group; chi-square tests of independence showed no evidence that treatment and control conditions differed on any student characteristic, both within and across LM/non-LM groups (all ps $> .05$).

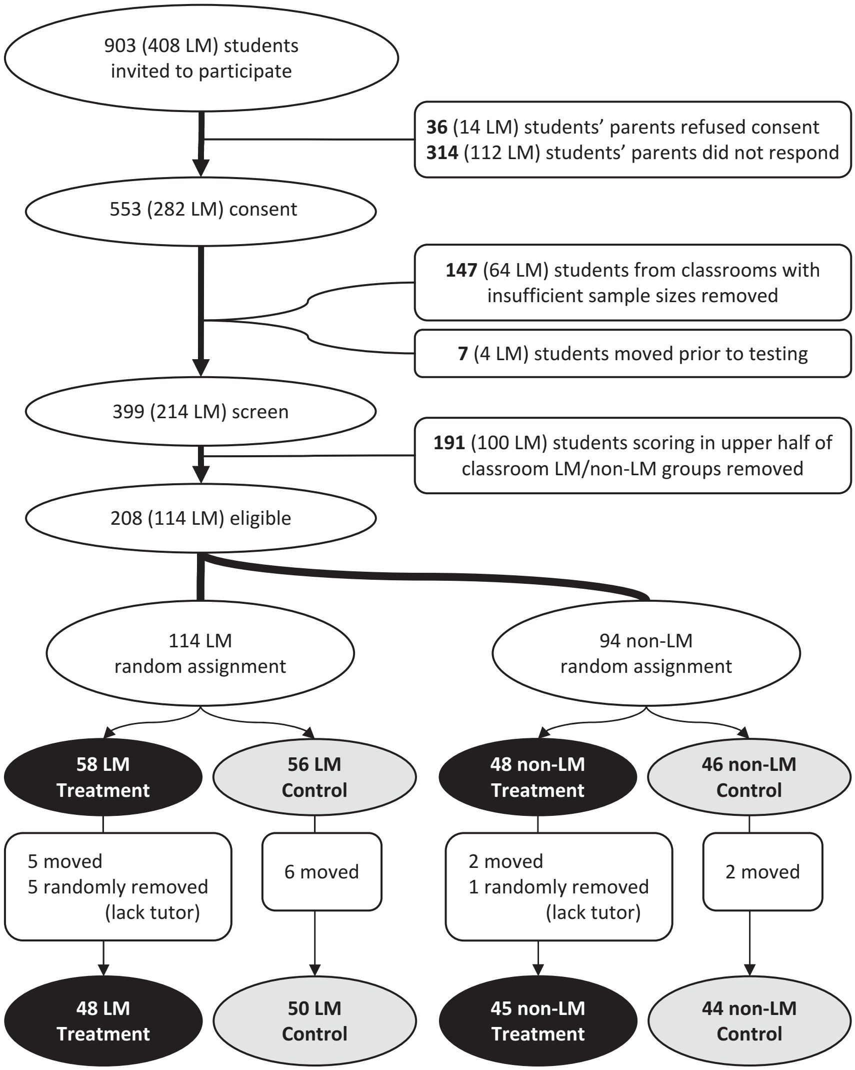


FIGURE 1 Flow chart of sample recruitment and assignment. LM = language minority.

Paraeducators. All paraeducator tutors (paraeducators) were recruited from their school communities based on their interest in working with children, prior tutoring and school volunteer experience, and scheduling flexibility. The 25 participating paraeducators were mostly nonminority (92%) and female (88%), had a median age of 45 to 54 years (ranging 18 to older than 55), and ranged

TABLE 1
Student Characteristics

	Non-LM (n = 89)				*LM (n = 98)*			
	Treatment (n = 48)		*Control (n = 45)*		*Treatment (n = 50)*		*Control (n = 44)*	
Characteristic	*N*	%	*N*	%	*N*	%	*N*	%
Male	24	53%	24	55%	34	71%	29	58%
EL	0	0%	0	0%	42	88%	46	92%
FRL	30	67%	28	64%	45	94%	43	86%
SPED	2	4%	2	5%	2	4%	3	6%
Minority	36	80%	34	77%	47	98%	49	98%
Asian	6	18%	11	22%	12	33%	8	17%
Black	20	59%	17	35%	5	14%	6	13%
Hispanic	5	15%	4	8%	29	81%	33	70%
Mixed/Other	5	15%	2	4%	1	3%	2	4%

Note. Chi-square tests of independence used to compare categorical frequencies between conditions across and within LM groups. All *p*s > .05 except EL services. LM = language minority; FRL = eligible for free or reduced lunch; EL = receives English learner services; SPED = receives special education services.

in educational level from high school diploma to master's degree, with a median level of a bachelor's degree (36%). Prior to the study, paraeducator tutoring experience ranged from 0 (24%) to 10 or more years (16%), with an average of 3.64 years ($SD = 4.54$). Most paraeducators (61%) had at least 1 year previous experience working with early grade levels (K-2; $M = 3.00$ years, $SD = 3.97$, range = 0–15 years). All paraeducators were hired as district employees. The assignment of students to tutors was determined by a combination of classroom scheduling, paraeducator availability, and the number of eligible students within classrooms within sites.

Paraeducator training. Researchers provided an initial 2-hr training session to describe each lesson activity and model paraeducator/student behaviors, errors, and error correction strategies. Trainees were paired together to practice each activity, whereas trainers provided feedback and responded to questions. Follow-up training was provided as needed throughout the intervention, with added coaching for paraeducators with less experience and/or low initial intervention fidelity ratings. Less experienced tutors received from 0.5 to 3.0 hr of added coaching during the intervention, averaging 1 hr of additional on-site coaching.

Intervention

Students assigned to treatment received 30 min of individual tutoring in English for 4 days per week, every week, between fall pretest and spring posttest

(November–May). Tutors were provided with a set of 108 scripted lessons, and tutors and students worked from these lesson pages. Each tutoring lesson included four to eight short components and was matched to decodable storybooks for oral reading practice. The five main lesson components are summarized as follows (for further details, see Jenkins et al., 2004; Vadasy & Sanders, 2008b).

- *Letter-sound correspondences.* Paraeducators introduced individual letter sounds, including digraphs. An average of one new letter sound was introduced in each lesson, with ongoing cumulative review of previously introduced letters.
- *Phoneme decoding*. Paraeducators modeled and introduced a phoneme decoding strategy, beginning with short decodable two- and three-letter words. Students were expected to orally blend the phonemes and pronounce each word.
- *Irregular words*. Paraeducators introduced and students practiced reading high-frequency irregular words before the words were scheduled to appear in the text reading. Students read, spelled, and reread each word.
- *Spelling*. Instruction and practice in spelling was integrated into several lesson components: letter sounds, phoneme decoding, and irregular words.
- *Oral reading practice*. The last 15 min of each tutoring session was allocated for oral reading practice in decodable texts. Paraeducators chose a reading method that matched each student's reading skills: independent reading, partner reading, or echo reading (most students read independently with tutor scaffolding). Paraeducators used scaffolding and correction procedures which emphasized immediate corrective feedback. Texts were drawn primarily from the *Bob Books* (Maslen, 2003), and students read each book at least three times.

Instructional scaffolding. Research staff provided ongoing coaching and modeling of appropriate scaffolding to help paraeducators provide the type of support at-risk students often require to accomplish phonemic segmenting, decoding, and encoding tasks (Foorman & Torgesen, 2001; Juel, 1996). Tutors who worked with LM students were instructed to adjust instruction to provide limited incidental vocabulary instruction that did not compromise intensity of the phonics instruction.

Intervention coverage. Mastery tests were employed to place students into the appropriate intervention lessons: Most (89%) were placed within the first 10 lessons, with the remaining students placed into the first 30 lessons (range = Lesson 1–31, $M = 3.97$, $SD = 6.85$). Tutors recorded daily student attendance (tutoring sessions) and lesson coverage. By the end of intervention, treatment students received on average 66.30 tutoring sessions ($M = 33.15$ hr of tutoring, $SD = 2.90$ hr), and completed 66.03 lessons ($SD = 15.76$). There were no significant differences between LM and non-LM students in lesson completion, tutoring attendance, or lesson-per-session coverage rate (all t-test p values $> .05$).

Treatment fidelity. Fidelity observations involved a 5-point rating scale ranging from 1 (*never implements correctly*) to 5 (*always implements correctly*) for each of the instructional components. After training but prior to field observations, researchers viewed six videotaped tutoring sessions of paraeducators implementing instruction with students. To determine interrater reliability, we calculated the internal consistency of the observers' mean implementation ratings for the videotaped sessions (using observers' ratings as items and each videotape as subjects): Cronbach's alpha was .97. After establishing reliability, researchers conducted a total of 240 observations over the course of the intervention, averaging 9.60 observations per tutor. Fidelity ratings across the 25 tutors' means averaged at 4.49 ($SD = 0.36$, range = 3.74–4.98).

Classroom Literacy Instruction Observations

In an effort to characterize typical classroom instruction for students in the study, particularly for control students, we conducted three formal observations of classroom literacy instruction blocks (approximately 2.5 months apart in November, February, and May).

We used an adapted version of the Instructional Content Emphasis–Revised (Edmonds & Briggs, 2003) to measure time afforded to dimensions of classroom literacy instruction. Seven certificated teacher observers were trained by the first author, and observer reliabilities were established through coding of sample videotapes of kindergarten and first-grade literacy instruction. Estimated reliabilities (using Cronbach's alpha) ranged from .74 (phonological awareness) to .99 (phonics/word study) for content categories ($Mdn = .95$, $M = .93$), and .99 for all grouping arrangement categories. Averaged across three observations, 29 first-grade teachers with participating students spent 87.02 min ($SD = 12.32$) on literacy block instruction, the majority of which was afforded to "other" (non-content) instruction (34–36%), text reading (17–22%), phonics/word study (10–20%), and comprehension (9–15%). Oral language was allocated on average 5% and vocabulary 3% of literacy block time. Teachers spent the smallest proportion of time on print concepts, phonological awareness, alphabetics, and fluency (< 1% on average). Finally, teachers used whole-class grouping (51–64% of their time) typically more than other grouping arrangements.

Student Assessments

Abilities hypothesized to contribute to or correlate with early word-reading skills were assessed at screening (September) or pretest (October), and included measures of receptive vocabulary, alphabetic knowledge, phonological awareness, word reading, and spelling. Posttesting took place at the end of the intervention (May) on alphabetic knowledge, phonological awareness, word reading, spelling, passage reading fluency, and comprehension. Tests were individually administered by trained testers who were unaware of student group assignment.

1. *Receptive (English) vocabulary* was measured at pretest only with the Peabody Picture Vocabulary Test–IIIA (Dunn & Dunn, 1997). Students selected a picture that best illustrates the meaning of an orally presented stimulus word. For this sample, internal consistency was .97 for all students and .96 for LM students.
2. *Alphabetic knowledge* was measured at pretest and posttest as the average of two naming measures: letter names and letter sounds correctly produced in 1 min. Both measures used all 26 letters of the alphabet twice (once in uppercase and once in lowercase). Letters were randomly sorted (uppercase and lowercase together) and presented on a single page in six rows in Comic Sans font, which allowed students to better differentiate between the lowercase letter L and uppercase letter I in particular. For the letter names task, students were asked to name as many letters on the page as they could. For the letter sounds task, students were asked to produce the sounds that each letter on the page represented. The number correctly named or produced, respectively, was divided by the number of seconds the students took to finish the 52 items. This was then multiplied by 60 to obtain letters correct per minute. For letter sounds, we considered hard consonants and soft vowels as correct. Internal consistency was .96 and .97 for letter names and sounds at pretest (.97 and .98 for LM students), and .92 and .89 at posttest (for LM students .94 and .91, respectively).
3. *Phonological awareness* was measured at pretest and posttest using a composite standard score of three subtests from the Comprehensive Test of Phonological Processing (Wagner et al., 1999): Blending Words, Elision, and Sound Matching. In the Blending Words subtest, the student was asked to listen to parts of words and blend them together to make a whole word. Internal consistency was .83 and .78 for pretest and posttest, respectively (for all students as well as LM students). During the Elision subtest, the student was asked to listen to the sounds in a spoken word and then to say the word without one or more of its sounds, creating a new word (e.g., "Say *tan*. Now say *tan* without saying /t/."). Internal consistency was .85 at pretest and .88 at posttest (for LM students, the respective values were .85 and .87). The Sound Matching subtest has two parts: In Part I, the tester said a word and asked the student to say, out of three choices, the word that started with the same sound as the initial word (e.g., "Which word starts with the same sound as *sock*? *Sun*, *cake*, or *bear*?"). Part II of this subtest asked the student to choose, out of three choices, the word that *ended* with the same sound as the initial word. Internal consistency was .90 at pretest (.86 for LM students), and .91 at posttest (.90 for LM students).

4. *Word reading* was measured at pretest and posttest using the Word Attack and Word Identification subtests from the Woodcock Reading Mastery Test–Revised/Normative Update (Woodcock, 1987–1998). The Word Attack subtest includes 45 nonwords that increase in difficulty. Internal consistency was .89 (.88 for LM students) and .94 (same for LM students) at pretest and posttest, respectively. The Word Identification subtest consists of 106 words that increase in difficulty. For this sample, internal consistency was .96 at pretest (.95 for LM students), and .96 at posttest (same for LM students).
5. *Spelling* was assessed at pretest and posttest using developmental raw scores derived from words correctly spelled on the Wide Range Achievement Test–Revised (Jastak & Wilkinson, 1984) Spelling subtest. This subtest requires the student to copy marks/symbols, print his or her name, and print a list of dictated words. Testing is discontinued after 10 consecutive incorrect responses. Similar to Fuchs et al. (2001), we applied the Tangel and Blachman (1992) developmental scoring rubric to all words attempted. This rubric allowed us to credit students for partial and less phonemically sophisticated responses. Items were scored from 0 (random string of letters) to 6 (entire word correctly spelled). Internal consistency for this sample was .96 at pretest and posttest (.96 and .95 for LM students for pretest and posttest, respectively).
6. *Passage reading fluency* was assessed at posttest only using the average words correctly read in one minute on two passages (Makar, 1996): one that we considered more decodable ("Mac and Tab") and one was considered less decodable ("Ben Bug"). Students read each passage aloud for 1 min. Words omitted, substituted, and hesitations of more than 3 s were considered errors. Words self-corrected within 3 s were scored as accurate. Internal consistencies were .99 (.98 for LM students) for both passages. The correlation between the two passages (alternate-form reliability) across students was .90 (.87 for LM students).
7. *Reading comprehension* was assessed at posttest only using the standard score of the Woodcock Reading Mastery Test–Revised/Normative Update Passage Comprehension subtest. The student was asked to silently read a short passage and then orally provide the missing key word. The student's raw score was the total number correct. Internal consistency was .92 (.90 for LM students).

Analysis Strategy

We adopted multilevel modeling as our primary analytic tool. In all of our analyses, we employed three-level models in which student scores (Level 1) were treated as nested within classrooms (Level 2, $n = 29$), which were in turn nested within schools (Level 3, $n = 11$). The t test of the slope for treatment (as well as the slopes for LM status and the interaction term) is similar to the classical t test, except that variance associated with nesting structures (i.e., between classrooms and schools) is explicitly estimated and accounted for in the predicted values and standard errors. For all analyses, treatment status and LM status were effect coded (i.e., $+1$ = treatment, -1 = control; $+1$ = LM, -1 = non-LM).

To answer our primary research questions (RQ1 and RQ2), we employed a three-level model that included treatment, LM status, and pretest word reading as predictors. We grand-mean centered word reading scores (i.e., z scores) for ease of results interpretation. Because there was no evidence supporting their exclusion, we included all interaction terms to help ensure model internal validity (functional form).

Finally, to test our auxiliary research question (RQ3), we adapted our posttest model for LM students only and replaced LM status with pretest receptive vocabulary as a predictor. Similar to word reading, we grand-mean centered vocabulary scores. We again included all interaction terms.

In all multilevel analyses, we used hierarchical linear modeling (Raudenbush, Bryk, & Congdon, 2004); SPSS (SPSS, Inc., 1989–2004) was used to compute classic statistics.

RESULTS

Pretest Models

Table 2 reports descriptive statistics by LM/non-LM group and experimental condition (zero-order correlations among all variables can be found in the appendix). Results from our pretest three-level multilevel models revealed no significant differences between the treatment and control groups; however, as expected (and as is evident in the descriptive statistics), there were significant differences between LM and non-LM students (a table of all model estimates are available from the first author upon request). Finally, we found significant pretest differences between schools on each measure, and between classrooms on receptive vocabulary, alphabetics, and spelling (χ^2 test p values $< .05$).

Primary Research Questions: Treatment Efficacy and Moderators

To answer our first and second research questions, we tested the effects of treatment status, LM status, pretest word reading level, and corresponding interactions simultaneously on each posttest score. As shown in Table 3, results revealed positive treatment effects on all but one outcome (Cohen's d calculated and reported in Table 2 for simple treatment effects on each posttest within LM/non-LM group). Holding all other variables constant, treatment students averaged 9.42 more letters correct per minute than controls, 5.12 more standard score points on word reading, 15.94 more raw points on developmental spelling, 12.52 more words correct

TABLE 2
Student Assessment Simple Means and Standard Deviations

	Non-LM (n = 89)								
	Treatment (n = 45)				*Control (n = 44)*				
	Pretest		*Posttest*		*Pretest*		*Posttest*		
Measure	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*d*
Recept Vocab	93.51	12.93			94.50	13.15			
Alphabetics	31.68	9.68	60.75	16.15	30.70	12.09	49.00	12.02	.82
Phono Aware	91.04	10.97	94.62	10.46	91.27	12.80	95.20	10.95	−.05
Word Reading	103.98	12.33	110.62	12.00	104.01	10.80	104.57	11.47	.52
Spelling	58.20	29.71	116.00	34.10	65.05	35.81	95.91	36.94	.57
Fluency			58.99	32.21			38.95	25.10	.69
Compreh			100.49	11.61			97.32	10.24	.29
	LM (n = 98)								
	Treatment (n = 48)				*Control (n = 50)*				
	Pretest		*Posttest*		*Pretest*		*Posttest*		
Measure	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*d*
Recept Vocab	73.90	16.24			74.06	13.44			
Alphabetics	26.43	14.14	54.34	18.20	27.25	11.39	48.69	15.74	.33
Phono Aware	84.65	12.92	88.90	12.60	82.94	10.11	86.78	10.25	.18
Word Reading	99.29	13.07	105.93	13.08	100.53	10.90	103.67	11.74	.18
Spelling	54.63	39.58	106.02	35.60	55.58	32.86	98.38	33.95	.22
Fluency			43.13	22.32			39.07	22.46	.18
Compreh			95.06	10.97			93.94	10.46	.10

Note. $N = 187$ students from 29 classrooms and 11 schools. Cohen's d reported is computed as the difference between posttest means divided by the pooled standard deviation. Receptive Vocab = Peabody Picture Vocabulary Test–IIIA standard score; Alphabetics = mean of letter names correct per minute and letter sounds correct per minute; Phono Aware = Comprehensive Test of Phonological Processing Phonological Awareness standard score; Word Reading = mean of Woodcock Reading Mastery Test Revised/Normative Update (WRMT–R/NU) Word Identification and Word Attack subtest standard scores; Spelling = Wide Range Achievement Test-Revised Spelling subtest developmental score of words dictated; Fluency = mean words correctly read per minute on two grade-level text passages; Comprehension = WRMT–R/NU Passage Comprehension standard score; LM = language minority.

per minute on passage reading fluency, and 2.80 more standard score points on comprehension. Not surprisingly, LM students performed lower than their non-LM peers at the end of first grade, although only for phonological awareness was

TABLE 3
Three-Level Posttest Models: Treatment Efficacy and Treatment Moderators

	Alphabetics		*Phono Aware*		*Word Reading*		*Spelling*		*Fluency*		*Compreh*	
Fixed Effects	*Coeff*	*SE*	*Coeff*	*SE*	*Coeff*	*SE*	*Coeff*	*SE*	*Coeff*	*SE*	*Coeff*	*SE*
M Posttest[a]	53.22	1.80***	91.15	0.97***	106.20	0.90***	103.46	2.26***	44.39	1.62***	96.64	0.65***
Treat[b]	4.71	0.87***	0.84	0.45	2.56	0.57***	7.97	1.92***	6.26	1.41***	1.40	0.48**
LM[b]	−1.89	1.22	−2.15	0.92*	0.05	0.67	1.63	2.48	−1.22	1.18	−0.68	0.50
Treat × LM[b]	−1.68	0.93	0.92	0.40*	−0.79	0.46	−3.09	1.66	−3.60	1.00***	−0.31	0.48
WR[b]	4.80	1.16***	6.93	0.60***	9.76	0.42***	21.19	1.91***	16.26	0.99***	8.23	0.52***
Treat × WR[b]	−0.71	0.79	0.25	0.65	−0.59	0.36	−2.73	1.77	1.19	1.08	−0.34	0.49
LM × WR[b]	−0.27	0.95	−1.02	0.84	−0.17	0.41	−2.07	2.48	−4.03	1.24**	−0.68	0.51
Treat × LM × WR[b]	1.28	1.08	1.26	0.51*	0.35	0.48	2.61	1.26*	−0.89	1.29	0.34	0.51
Random Effects	*Variance*		*Variance*		*Variance*		*Variance*		*Variance*		*Variance*	
Classrooms[c]	0.00		1.90		5.14*		0.12		33.06*		6.25*	
Schools[d]	24.79***		2.94		4.68*		6.94		0.16		0.01	
Residual	194.42		66.72		39.96		768.02		341.78		40.66	

Note. $N = 187$ students from 29 classrooms and 11 schools. Alphabetics = mean of letter names correct per minute and letter sounds correct per minute; Phono Aware = Comprehensive Test of Phonological Processing Phonological Awareness standard score; Word Reading = mean of Woodcock Reading Mastery Test–Revised/Normative Update (WRMT–R/NU) Word Identification and Word Attack subtest standard scores; Spelling = Wide Range Achievement Test–Revised Spelling subtest developmental score of words dictated; Fluency = mean words correctly read per minute on two grade-level text passages; Compreh = WRMT-R/NU Passage Comprehension standard score; Treat = treatment status effect-coded +1 = treatment, −1 = control; LM = language minority status effect-coded +1 = LM, −1 = non-LM; WR = pretest word reading grand-mean centered.

[a] t test $df = 10$. [b] t test $df = 179$. [c] Chi-square test $df = 18$. [d] Chi-square test $df = 10$.

$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

the difference between LM students and non-LM students significant. Further, as expected, pretest word reading was significantly predictive of all posttests.

Of more interest, results reveal that treatment response was only moderated by LM status for posttest phonological awareness (LM students exhibit greater treatment effects than non-LM students) and passage reading fluency (LM students have lower treatment effects than non-LM students). Further, pretest word reading did not generally moderate treatment response; however, this was qualified by significant three-way interactions detected for phonological awareness and spelling. Figure 2 displays the model-predicted posttest means for these two outcomes by group and pretest word reading level (lower = 1 *SD* below average; higher = 1 *SD* above average). For phonological awareness, higher pretest word reading was associated with a stronger treatment effect for LM students but not for non-LM students. For spelling, in contrast, lower pretest word reading was associated with a stronger treatment effect for non-LM students but not for LM students. Finally, although not illustrated, results also show that pretest word reading negatively interacted with LM status on posttest passage reading fluency: the relationship between pretest and posttest was stronger for non-LM students compared to LM students.

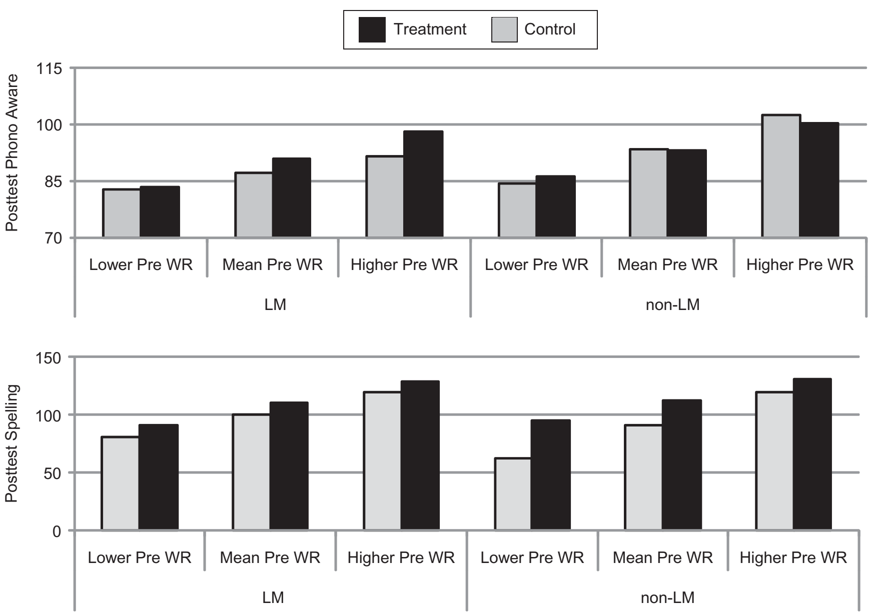


FIGURE 2 Model-estimated means for posttests with significant three-way interactions among treatment, language minority (LM) status, and pretest word reading (pre WR).

Auxiliary Research Question: Vocabulary as Treatment Moderator for LM Students

Finally, we examined the possibility that students require some level of proficiency in English vocabulary knowledge before they can truly benefit from a reading intervention. Although we cannot tease apart whether LM status or vocabulary knowledge moderates treatment effect (we had too few non-LM language-impaired students in our sample), we examined whether LM students with higher levels of vocabulary benefit differentially from treatment.

Model results (Table 4) reveal that LM treatment students outperformed LM controls on posttest phonological awareness, word reading, and spelling. Not surprisingly, pretest vocabulary was significantly uniquely predictive of LM students' phonological awareness and comprehension posttests, and pretest word reading was significantly associated with all posttests except alphabetics. Results show that LM students' responsiveness to treatment was moderated by pretest vocabulary for posttest word reading (two-way interaction), spelling (two- and three-way interactions), and comprehension (three-way interaction; see Figure 3).

DISCUSSION

This study examined the benefits of individual supplemental phonics-based instruction for relatively low-skilled LM and non-LM first graders. Within each of the 29 participating classrooms, the bottom halves of LM and non-LM students were randomly assigned to either treatment (phonics-based supplemental tutoring) or control (classroom instruction only) conditions. Inclusion of both LM and non-LM treatment and control conditions allowed us to adequately test whether treatment effects differed for LM students compared with non-LM students. It is important to note that use of a relative-risk study eligibility criterion within classrooms, rather than an absolute cut-value risk criterion, meant that students entered this study with higher and more varied word reading ability than what is typical for intervention studies. This variability allowed us to test whether baseline word reading ability (as a risk indicator) was predictive of treatment response. Finally, for LM students, we also tested whether baseline receptive vocabulary moderated responsiveness to intervention.

Our finding of significant positive treatment effects on all but one of our outcome measures (across LM and non-LM students) replicates earlier results on the effectiveness of paraeducator-implemented, supplemental phonics-based instruction for at-risk kindergarteners (Vadasy & Sanders, 2008a; Vadasy, Sanders, & Peyton, 2006a), first graders (Jenkins et al., 2004; Vadasy et al., 1997a, 1997b), and children in Grades 2 and 3 (Vadasy, Sanders, & Peyton, 2006b; Vadasy,

TABLE 4
Three-Level Posttest Models: Treatment Moderators for LM Students

	Alphabetics		*Phono Aware*		*Word Reading*		*Spelling*		*Fluency*		*Compreh*	
Fixed Effects	*Coeff*	*SE*	*Coeff*	*SE*	*Coeff*	*SE*	*Coeff*	*SE*	*Coeff*	*SE*	*Coeff*	*SE*
Mean Posttest[a]	51.17	2.36***	87.64	1.08***	104.58	1.27***	102.57	2.83***	41.10	2.02***	94.48	0.88***
Treat[b]	2.69	1.61	1.81	0.45***	1.87	0.68**	6.71	3.27*	2.60	1.70	1.24	0.80
VOC[b]	1.35	1.88	2.72	0.69***	1.25	0.65	3.13	3.09	−0.28	2.50	2.45	0.73**
Treat × VOC[b]	1.95	1.18	1.08	0.64	1.07	0.52*	4.47	1.24***	2.02	1.50	1.13	0.92
WR[b]	3.93	2.05	5.51	0.87***	9.39	0.55***	18.22	3.18***	12.82	2.32***	6.96	0.55***
Treat × WR[b]	0.47	1.33	0.78	0.79	−0.60	0.46	−2.80	2.43	−0.15	1.77	−0.72	0.44
WR × VOC[b]	−0.96	1.48	1.31	0.64*	0.21	0.33	−1.09	3.36	0.60	1.11	−0.40	0.58
Treat × WR × VOC[b]	0.51	1.85	−1.34	0.71	−0.59	0.37	−5.87	2.74*	0.23	1.12	−0.84	0.39*
Random Effects	*Variance*		*Variance*		*Variance*		*Variance*		*Variance*		*Variance*	
Classrooms[c]	0.28		4.90		0.39		0.74		29.06*		0.03	
Schools[d]	35.76**		2.79		11.02***		8.76		0.09		3.79*	
Residual	215.24		69.18		42.08		717.27		287.73		39.94	

Note. $N = 98$ students from 29 classrooms and 11 schools. Alphabetics = mean of letter names correct per minute and letter sounds correct per minute; Phono Aware = Comprehensive Test of Phonological Processing Phonological Awareness standard score; Word Reading = mean of Woodcock Reading Mastery Test–Revised/Normative Update (WRMT–R/NU) Word Identification and Word Attack subtest standard scores; Spelling = Wide Range Achievement Test–Revised Spelling subtest developmental score of words dictated; Fluency = mean words correctly read per minute on two grade-level text passages; Compreh = WRMT-R/NU Passage Comprehension standard score; Treat = treatment status effect-coded +1 = treatment, −1 = control; LM = language minority status effect-coded +1 = LM, −1 = non-LM; VOC = pretest receptive vocabulary grand-mean centered; WR = pretest word reading grand-mean centered.

[a]t test $df = 10$. [b]t test $df = 90$. [c]Chi-square test $df = 18$. [d]Chi-square test $df = 10$.

*$p < .05$. **$p < .01$. ***$p < .001$.

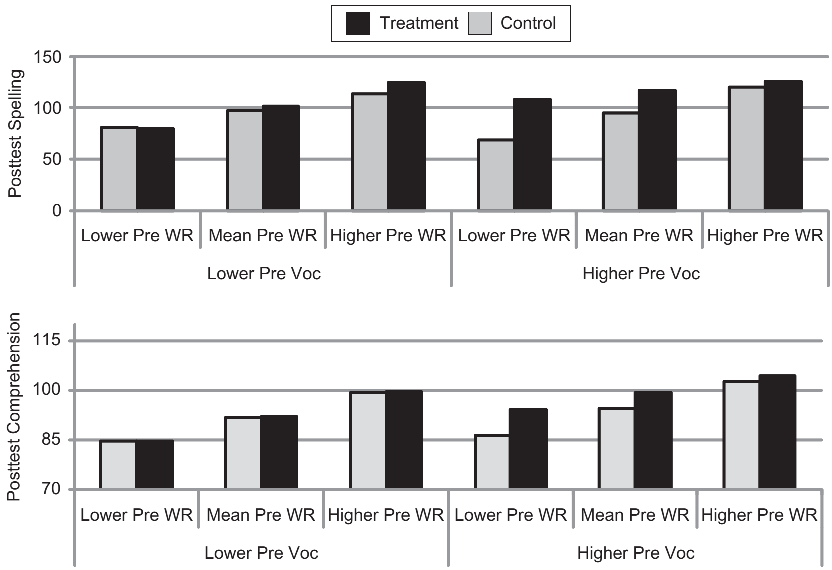


FIGURE 3 Model-estimated means for language minority students only for posttests with significant three-way interactions among treatment, pretest word reading (pre WR), and pretest vocabulary (pre VOC).

Sanders, & Tudor, 2007). Although pretest word reading ability was positively associated with posttest scores, treatment effects were generally not moderated by pretest word reading. The lack of systematic interactions between word reading and treatment provides evidence that both students with *relatively* poor (compared to their classroom LM/non-LM peers) and good word reading skills can benefit from individual supplemental phonics-based instruction.

Despite the positive findings overall for treatment, results showed that LM students tended to exhibit lower treatment effects than non-LM students, although significantly lower only on passage reading fluency. Simple effect sizes (Cohen's *d*) reported in Table 2 for LM students are approximately one third of the size of effects for non-LM students. Moreover, in our auxiliary analyses with only the LM students, we found that treatment response on word reading and spelling was positively predicted by pretest vocabulary. As illustrated in Figure 3, LM students with higher vocabulary (in our sample, 1 *SD* above average for LM students is approximately the 25th percentile) appear to experience greater treatment benefits compared with LM peers with less vocabulary knowledge who showed limited or no advantage from treatment. Word-level outcomes (word reading and spelling) can be regarded as near-transfer outcomes for a phonics intervention such as that tested in our study. Others have observed that vocabulary knowledge contributes unique variance to word reading in young children, possibly through semantic compensatory support for word reading (Ricketts, Nation, & Bishop, 2007). Higher vocabulary knowledge may also benefit LM students if larger vocabulary size leads to lexical restructuring and segmental word representation that facilitate phonological awareness and developmental spelling.

Instructional Implications

There is limited research to guide decisions for intervention planning for LM students who are at risk for reading difficulties. The results of the present study suggest that treatment effects for LM students across reading outcomes were small. Observations of typical classroom instruction suggest that classroom teachers in this study allocated up to 20% of instruction time for phonics/word-level skills. In classrooms where teachers provide less phonics instruction, treatment effects for LM students may have been higher.

On the other hand, we cannot overlook the extremely low levels of pretest receptive vocabulary for LM students relative to their pretest word reading scores and relative to the pretest vocabulary levels for non-LM students. Such limited vocabulary knowledge constrains comprehension as well as word reading development. Our research staff observed tutors implementing vocabulary instruction effectively when it was provided, suggesting that paraeducators are able to scaffold vocabulary learning for LM students with limited English skills. Findings suggest that at-risk LM students presenting with low levels of vocabulary warrant intervention with an oral language/vocabulary focus. In schools like the sites in this study where bilingual instruction is unevenly provided, it may be more valuable to provide at-risk LM students with supplemental intervention that has a greater emphasis on English oral language, vocabulary, and reading comprehension. A language-focussed supplemental intervention may better complement instruction for students like those in this study whose classroom phonics instruction is adequate to develop necessary decoding skills but who do not receive sufficient intensive vocabulary or language instruction to catch up with peers.

Limitations

The current study has some important limitations. First, our sampling design did not employ a typical cut-value for screening children into the study; instead, we used a relative-performance criterion in which students included in the study were those who performed in the bottom half of their own classrooms, within their LM/non-LM group. This is a liberal cutoff for an intervention study; however, it ensured adequate group sample sizes to test the interaction between treatment and LM status. The result of our sampling method meant that our sample was generally higher and more varied in early reading skills than would typically be identified for intervention. As such, findings from this study generalize to students who are performing relatively low compared to their own classroom peers. Nevertheless, we note that pretest word reading was not generally found to moderate treatment response on posttests, and therefore students with higher levels of word reading than are typically considered for intervention may benefit in a similar fashion to more at-risk peers. Second, our definition of LM status was based on parent self-report of language primarily spoken in the home rather than students' actual English language proficiency. Although there is precedent for using this definition in previous studies (e.g., Lipka & Siegel, 2007; Silverman & Hines, 2009), lack of a measure of students' English proficiency did not allow us to demonstrate the level of second-language proficiency necessary for adequate treatment response. Third, we did not have information on language proficiency levels in students' primary home language, or on early home literacy practices that might account for differences in L2 reading development. Indeed, the language diversity of this sample precluded obtaining home language proficiency levels. Fourth, our models do not account for bilingual services students received, which varied across schools in part based on the proportion of LM students enrolled. Nevertheless, we did account for between-school and between-classroom variation in the dependent measures in all of our models. Finally, the treatment effects sizes found in the current study apply to individual 1:1 tutored instruction, a costly delivery model for many schools to adopt. Future research should examine the benefits of small-group reading intervention for LM students.

Conclusion

The fact that this tutoring intervention was effectively implemented by paraeducators extends options available to schools seeking to augment first-grade classroom phonics instruction for lower skilled students. We have reported previously on the effectiveness of similar paraeducator-implemented instruction for L1 English-speaking students with skill levels somewhat lower than those for non-LM students in this study. A practical finding in this study is that paraeducators effectively implemented the intervention with LM students, and required minimal added support to instruct students with very limited English skills.

Findings indicate differential benefits for supplementing classroom reading instruction with phonics-based reading tutoring for low performing LM and non-LM first-grade students. Treatment effects were nonexistent for LM students for fluency and comprehension, the skills for which LM students may most need intervention to catch up and keep up with non-LM peers. It is possible that helping to consolidate word-level skills will have later benefits for LM students in fluency or comprehension, although this requires further study.

Compared to non-LM students' response to treatment, the limited response of LM students to phonics-based tutoring must be considered in light of the cost and logistics of providing supplemental instruction. In schools, like the research sites in this study, with a strong phonics emphasis in the core literacy program, paraeducators may be better utilized to supplement oral language and vocabulary instruction for LM students. Our findings for LM students suggest that schools should consider the adequacy of both word-level and language-level classroom resources available for their LM students in choosing targets for tutoring interventions.

REFERENCES

Allor, J., & McCathren, R. (2004). The efficacy of an early literacy tutoring program implemented by college students. *Learning Disabilities Research & Practice, 19*, 116–129.

August, D., & Hakuta, K. (1998). *Educating language minority children*. Washington, DC: National Academy Press.

August, D., & Shanahan, T. (2006). *Developing literacy in second-language learners: Report of the National Literacy Panel on Language-Minority Children and Youth*. Mahwah, NJ: Erlbaum.

Baker, S., Gersten, R., & Keating, T. (2000). When less may be more: A 2-year longitudinal evaluation of a volunteer tutoring program requiring minimal training. *Reading Research Quarterly, 35*, 494–519.

Betts, J., Bolt, S., Decker, D., Muyskens, P., & Marston, D. (2009). Examining the role of time and language type in reading development for English Language Learners. *Journal of School Psychology, 47*, 143–166.

Blachman, B. A., Ball, E., Black, R., & Tangel, D. (1994). Kindergarten teachers develop phoneme awareness in low-income, inner-city classrooms: Does it make a difference? *Reading and Writing: An Interdisciplinary Journal, 6*, 1–17.

Brown, K., Morris, D., & Fields, M. (2005). Intervention after Grade 1: Serving increased numbers of struggling students effectively. *Journal of Literacy Research, 37*, 61–94.

Canadian Education Statistics Council. (2007). *Education indicators in Canada: Report of the Pan-Canadian Education Indicators Program 2007*. Toronto, Ontario: Author.

Capps, R., Fix, M., Murray, J., Ost, J., Passel, J., & Herwantoro, S. (2005). *The new demography of America's schools: Immigration and the No Child Left Behind Act*. Washington, DC: The Urban Institute.

Center, Y., Wheldall, K., Freeman, L., Outhred, L., & McNaught, M. (1995). An evaluation of Reading Recovery. *Reading Research Quarterly, 30*, 240–263.

Chiappe, P., & Siegel, L. S. (2006). A longitudinal study of reading development of Canadian children from diverse language backgrounds. *The Elementary School Journal, 107*, 135–152.

Denton, C. A., Anthony, J. L., Parker, R., & Hasbrouck, J. E. (2004). Effects of two tutoring programs on the English reading development of Spanish–English bilingual students. *The Elementary School Journal, 104*, 289–305.

Droop, M., & Verhoeven, L. (2003). Language proficiency and reading ability in first- and second-language learners. *Reading Research Quarterly, 38*, 78–103.

Dunn, L. M, & Dunn, L. M. (1997). *Peabody Picture Vocabulary Test–IIIA* (3rd ed.). Circle Pines, MN: American Guidance Service.

Edmonds, M., & Briggs, K. L. (2003). The Instructional Content Emphasis Instrument: Observations of reading instruction. In S. Vaughn & K. L. Briggs (Eds.), *Reading in the classroom: Systems for the observation of teaching and learning* (pp. 31–47). Baltimore, MD: Brookes.

Ehri, L., Dreyer, L., Flugman, B., & Gross, A. (2007). Reading Rescue: An effective tutoring intervention model for language-minority students who are struggling readers in first grade. *American Educational Research Journal, 44*, 414–448.

Ehri, L. C., Nunes, S. R., Stahl, S. A., & Willows, D. M. (2001). Systematic phonics instruction helps students learn to read: Evidence from the National Reading Panel's meta-analysis. *Review of Educational Research, 71*, 393–447.

Entwisle, D. R., Alexander, K. L., & Olson, L. S. (1997). *Children, schools, and inequality*. Boulder, CO: Westview.

Eurydice. (2004). *Integrating immigrant children into schools in Europe*. Brussels, Belgium: Eurydice European Unit.

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

Fry, R. (2008). *The role of schools in the English Language Learner achievement gap*. Washington, DC: Pew Hispanic Center.

Fuchs, D., Fuchs, L. S., Thompson, A., Al Otaiba, S., Yen, L., Yang, N. J., . . . O'Connor, R. E. (2001). Is reading important in reading-readiness programs? A randomized field trial with teachers as program implementers. *Journal of Educational Psychology, 93*, 251–267.

Gelzheiser, L. M. (2005). Maximizing student progress in one-to-one programs: Contributions of texts, volunteer experience, and student characteristics. *Exceptionality, 13*, 229–243.

Gottardo, A., & Mueller, J. (2009). Are first- and second-language factors related in predicting second-language reading comprehension? A study of Spanish-speaking children acquiring English as a second language from first to second grade. *Journal of Educational Psychology, 101*, 330–344.

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education, 7*, 6–10.

Gunn, B., Biglan, A., Smolkowski, K., & Ary, D. (2000). The efficacy of supplemental instruction in decoding skills for Hispanic and non-Hispanic students in early elementary school. *The Journal of Special Education, 34*, 90–103.

Hatcher, P. J., Hulme, C., & Ellis, A. W. (1994). Ameliorating early reading failure by integrating the teaching of reading and phonological skills: The phonological linkage hypothesis. *Child Development, 65*, 41–57.

Hatcher, P. J., Hulme, C., Miles, J. N. V., Carroll, J. M., Hatcher, J., Gibbs, S., . . . Snowling, M. J. (2006). Efficacy of small group reading intervention for beginning readers with reading-delay: A randomized controlled trial. *Journal of Child Psychology and Psychiatry, 47*, 820–827.

Heath, A. F., Rothon, C., & Kilpi, E. (2008). The second generation in Western Europe: Education, unemployment, and occupational attainment. *The Annual Review of Sociology, 34*, 211–235.

Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing: An Interdisciplinary Journal, 2*, 127–160.

Jastak, S., & Wilkinson, G. S. (1984). *The Wide Range Achievement Test–Revised*. Wilmington, DE: Jastak Associates.

Jenkins, J. R., Hudson, R. F., & Johnson, E. S. (2007). Screening for at-risk readers in a response to intervention framework. *School Psychology Review, 36*, 582–600.

Jenkins, J. R., Peyton, J. A., Sanders, E. A., & Vadasy, P. F. (2004). Effects of reading decodable texts in supplemental first-grade tutoring. *Scientific Studies of Reading, 8*, 58–86.

Juel, C. (1996). What makes literacy tutoring effective? *Reading Research Quarterly, 31*, 268–289.

Kieffer, M. J. (2008). Catching up or falling behind? Initial English proficiency, concentrated poverty, and the reading growth of language minority learners in the United States. *Journal of Educational Psychology, 100*, 851–868.

Lesaux, N. K. (2006). Building consensus: Future directions for research on English Language Learners at risk for learning disabilities. *Teachers College Record, 11*, 2406–2438.

Lesaux, N. K., & Geva, E. (2006). Synthesis: Development of literacy in language-minority students. In D. August & T. Shanahan (Eds.), *Developing literacy in second-language learners: Report of the National Literacy Panel on Language-Minority Children and Youth* (pp. 53–74). Mahwah, NJ: Erlbaum.

Lesaux, N. K., & Siegel, L. S. (2003). The development of reading in children who speak English as a second language. *Developmental Psychology, 39*, 1005–1019.

Lipka, O., & Siegel, L. S. (2007). The development of reading skills in children with English as a second language. *Scientific Studies of Reading, 11*, 105–131.

Makar, B. (1996). *Primary Phonics–1*. Cambridge, MA: Educators Publishing Service.

Maslen, B. L. (2003). *Bob books*. New York, NY: Scholastic.

Mathes, P. G., Denton, C. A., Fletcher, J. M., Anthony, J. L., Francis, D. J., & Schatschneider, C. (2005). The effects of theoretically different instruction and student characteristics on the skills of struggling readers. *Reading Research Quarterly, 40*, 148–182.

O'Connor, R. E. (2000). Increasing the intensity of intervention in kindergarten and first grade. *Learning Disabilities Research and Practice, 15*, 43–54.

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.

O'Connor, R. E., & Jenkins, J. R. (1999). The prediction of reading disabilities in kindergarten and first grade. *Scientific Studies of Reading, 3*, 159–197.

Phillips, M., Crouse, J., & Ralph, J. (1998). Does the black/white test score gap widen after children enter school? In C. Jencks & M. Phillips (Eds.), *The black/white Test Score Gap* (pp. 229–272). Washington, DC: Brookings.

Planty, M., Hussar, W., Snyder, T., Kena, G., Kewal Ramani, A., Kemp, J., . . . Dinkes, R. (2009). *The Condition of Education 2009* (NCES 2009–081). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Proctor, C. P., August, D., Carlo, M. S., & Snow, C. (2006). The intriguing role of Spanish language vocabulary knowledge in predicting English reading comprehension. *Journal of Educational Psychology, 98*, 159–169.

Raudenbush, S. W., Bryk, A. S., & Congdon, R. T. (2004). *HLM for Windows 6.0*. Lincolnwood, IL: Scientific Software International, Inc.

Ricketts, J., Nation, K., & Bishop, D.V.M. (2007). Vocabulary is important for some, but not all reading skills. *Scientific Studies of Reading, 11*, 235–257.

Schneider, W., Roth, E., & Ennemoser, M. (2000). Training phonological skills and letter knowledge in children at risk for dyslexia: A comparison of three kindergarten intervention programs. *Journal of Educational Psychology, 92*, 284–295.

Shanahan, T., & Beck, I. L. (2006). Effective literacy teaching for English-Language Learners. In D. August & T. Shanahan (Eds.), *Developing literacy in second-language learners: Report of the National Literacy Panel on Language-Minority Children and Youth* (pp. 415–488). Mahwah, NJ: Erlbaum.

Silverman, R., & Hines, S. (2009). The effects of multimedia-enhanced instruction on the vocabulary of English-language learners and non-English-language learners in pre-kindergarten through second grade. *Journal of Educational Psychology, 101*, 305–314.

Slavin, R. E., Lake, C., Davis, S., & Madden, N. A. (2009). *Effective programs for struggling readers: A best-evidence synthesis*. Baltimore, MD: Johns Hopkins University/University of York.

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

SPSS, Inc. (1989–2004). *Statistical Package for the Social Sciences 13.0*. New York, NY: McGraw-Hill.

Tangel, D. M., & Blachman, B. A. (1992). Effect of phoneme awareness instruction on kindergarten children's invented spelling. *Journal of Reading Behavior, 24*, 233–261.

Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (1997). Prevention and remediation of severe reading disabilities: Keeping the end in mind. *Scientific Studies of Reading, 1*, 217–234.

Torgesen, J. K., Wagner, R. K., Rashotte, C. A., Herron, J., & Lindamood, P. (2009). *Computer assisted instruction to prevent early reading difficulties in students at risk for dyslexia: Outcomes from two instructional approaches*. Tallahassee, FL: Florida State University.

University of California Linguistic Minority Research Institute. (2007, Winter). Lagging behind: Linguistic minorities' educational progress during elementary school. *Newsletter, 16*, 1–3.

U.S. Department of Education, Institute of Education Sciences. (2007, July 30). *What Works Clearinghouse: English Language Learners*. Retrieved from http://ies.ed.gov/ncee/wwc/pdf/ELL_TR_07_30_07.pdf

Vadasy, P. F., Jenkins, J. R., Antil, L. R., Wayne, S. K., & O'Connor, R. E. (1997a). Community-based early reading intervention for at-risk first graders. *Learning Disabilities: Research and Practice, 12*, 29–39.

Vadasy, P. F., Jenkins, J. R., Antil, L. R., Wayne, S. K., & O'Connor, R. E. (1997b). The effectiveness of one-to-one tutoring by community tutors for at-risk beginning readers. *Learning Disability Quarterly, 20*, 126–139.

Vadasy, P. F., & Sanders, E. A. (2008a). Code-oriented instruction for kindergarten students at risk for reading difficulties: A replication and comparison of instructional grouping. *Reading and Writing: An Interdisciplinary Journal, 21*, 929–963.

Vadasy, P. F., & Sanders, E. A. (2008b). Individual tutoring for struggling readers: Moving research to scale with interventions implemented by paraeducators. In G. A. Reid, A. Fawcett, F. Manis, & L. Seigel (Eds.), *The SAGE handbook of dyslexia* (pp. 337–355). Thousand Oaks, CA: Sage.

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.

Vadasy, P. F., Sanders, E. A., & Peyton, J. A. (2006a). 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.

Vadasy, P. F., Sanders, E. A., & Peyton, J. A. (2006b). Paraeducator-supplemented instruction in structural analysis with text reading practice for second and third graders at risk for reading problems. *Remedial and Special Education, 27*, 365–378.

Vadasy, P. F., Sanders, E. A., & Tudor, S. (2007). Effectiveness of paraeducator-supplemented individual instruction: Beyond basic decoding skills. *Journal of Learning Disabilities, 40*, 508–525.

Vellutino, F. R., Scanlon, D. M., Sipay, E. R., Small, S. G., Pratt, A., Chen, R., 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.

Vellutino, F. R., Tunmer, W. F., Jaccard, J. J., & Chen, R. (2007). Components of reading ability: Multivariate evidence for a convergent skill model of reading development. *Scientific Studies of Reading, 11*, 3–32.

Verhoeven, L. (1990). Acquisition of reading in a second language. *Reading Research Quarterly, 25*, 90–114.

Wagner, R., Torgesen, J. K., & Rashotte, C. A. (1999). *Comprehensive Test of Phonological Processing*. Austin, TX: Pro-Ed.

Watson, T., & Hempenstall, K. (2008). Effects of a computer based beginning reading program on young children. *Australasian Journal of Educational Technology, 24*, 258–274.

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

Woodcock, R. (1987–1998). *Woodcock Reading Mastery Test–Revised/Normative Update*. Circle Pines, MN: American Guidance Service.

APPENDIX
Zero-Order Correlations

Variable	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*
Conditions												
1. Treatment status	—											
2. LM status	−.02	—										
Pretests												
3. Recept Vocab	−.01	−.58*	—									
4. Alphabetics	.01	−.18*	.30*	—								
5. Phono Aware	.04	−.30*	.62*	.45*	—							
6. Word Reading	−.03	−.17*	.36*	.54*	.66*	—						
7. Spelling	−.05	−.09	.31*	.59*	.54*	.66*	—					
Posttests												
8. Alphabetics	.26*	−.11	.12	.47*	.27*	.30*	.35*	—				
9. Phono Aware	.04	−.31*	.49*	.44*	.69*	.64*	.58*	.29*	—			
10. Word Reading	.17*	−.12	.31*	.46*	.58*	.79*	.57*	.48*	.65*	—		
11. Spelling	.19*	−.06	.20*	.41*	.49*	.57*	.61*	.50*	.61*	.68*	—	
12. Fluency	.22*	−.15*	.22*	.50*	.45*	.62*	.55*	.59*	.46*	.69*	.62*	—
13. Compreh	.10	−.20	.48*	.43*	.66*	.77*	.55*	.38*	.70*	.85*	.62*	.68*

Note. $N = 187$ students from 29 classrooms and 11 schools; Treatment (1 = treatment, 0 = control); LM = language minority (1 = LM, 0 = non-LM); Recept Vocab = Peabody Picture Vocabulary Test–IIIA standard score; Alphabetics = mean of letters correct per minute; Phono Aware = Comprehensive Test of Phonological Processing Phonological Awareness standard score; Word Reading = mean of Woodcock Reading Mastery Test Revised/Normative Update Word Identification and Word Attack subtest standard scores; Spelling = Wide Range Achievement Test–Revised Spelling subtest developmental score of words dictated; Fluency = passage reading fluency defined as mean words correctly read per minute on two grade-level text passages; Compreh = Woodcock Reading Mastery Test–Revised/Normative Update Passage Comprehension standard score. Pearson's r reported.

*$p < .05$.