Effectiveness of an English Intervention for First-Grade English Language Learners at Risk for Reading Problems

Sharon Vaughn
University of Texas at Austin

Patricia Mathes
Southern Methodist University

Sylvia Linan-Thompson
University of Texas at Austin

Paul Cirino
University of Houston

Coleen Carlson
University of Houston

Sharolyn Pollard-Durodola
Texas A & M University

Elsa Cardenas-Hagan
University of Houston

David Francis
University of Houston

The Elementary School Journal
Volume 107, Number 2

0013-5984/2006/10702-0003$05.00

Abstract

A first-grade reading and language development intervention for English language learners (Spanish/English) at risk for reading difficulties was examined. The intervention was conducted in the same language as students' core reading instruction (English). Two hundred sixteen first-grade students from 14 classrooms in 4 schools from 2 districts were screened in both English and Spanish. Forty-eight students (22%) did not pass the screening in both languages and were randomly assigned within schools to an intervention or contrast group; after 7 months, 41 students remained in the study. Intervention groups of 3 to 5 students met daily (50 minutes) and were provided systematic and explicit instruction in oral language and reading by trained bilingual reading intervention teachers. Students assigned to the contrast condition received their school's existing intervention for struggling readers. Intervention students significantly outperformed contrast students on multiple measures of English letter naming, phonological awareness and other language skills, and reading and academic achievement. Differences were less significant for Spanish measures of these domains, though the strongest effects favoring the intervention students were in the areas of phonological awareness and related reading skills.

Though the understanding of beginning reading instruction for students at risk for reading problems is incomplete, there is little question that researchers and educators have made great strides in designing effective interventions for monolingual English students at risk for reading problems (O'Connor, 2000; Torgesen, Mathes, & Grek, 2002). Syntheses of interventions (Ehri, Nunes, Stahl, & Willows, 2001; Kuhn & Stahl, 2003; National Reading Panel, 2000; Snow, Burns, & Griffin, 1998) report findings that improve confidence about many elements of reading instruction, and these find-

ings have also influenced policy (Fletcher & Lyon, 1998). Missing from the research on reading difficulties are numerous experimental studies with at-risk students who are English language learners (ELLs). In this article we (research team) report findings from a randomized, controlled trial providing a systematic and explicit intervention (reading instruction in addition to core reading) to enhance reading outcomes for ELLs (Spanish/English) at risk for reading problems and learning to read in English. This study was designed to enhance understanding of the extent to which ELLs acquiring literacy skills in English could progress and meet grade-level expectations in reading when provided an intervention program originally designed for monolingual English first graders at risk for reading problems and modified to be used as an ESL reading intervention.

Several issues related to early reading intervention with ELLs make the selection and implementation of an intervention even more complex than with monolingual students. Many ELLs who are at risk for reading problems in first grade are low in oracy skills in both English and Spanish (Geva, Yaghoub-Zadeh, & Schuster, 2000; Vaughn et al., 2006, in press). Little is known about the influence of these low oracy skills on the development of beginning literacy skills; some have questioned whether students with low language skills can profit from early reading instruction (Araujo, 2002; Edelsky, 1986; Fitzgerald & Noblit, 1999). There is evidence that development of language and literacy in some students can be enhanced through reading instruction and that reading instruction and reading development in English outpace their oral language development in English (Elley & Magubhai, 1983; Fitzgerald & Noblit, 1999; Geva & Wang, 2001; Linan-Thompson, Vaughn, Hickman-Davis, & Kouzekanani, 2003). Thus, the experimental study described in this article was conducted to assess the influence of an English intervention on the development of oracy and literacy skills for these students.

A related issue is the extent to which an English reading intervention for at-risk ELLs might improve oracy and literacy development in Spanish, though students are taught exclusively in English. Learning to read in a more difficult alphabetic language like English may provide students with foundational skills to enhance their reading skills in a more orthographically transparent language like Spanish.

Interventions with EL Learners

Reviews of research on reading instruction for ELLs have yielded few studies and little converging scientific evidence about best practices for students who require supplemental reading interventions (Fitzgerald, 1995a, 1995b; Gersten & Baker, 2000). Though research on the effectiveness of reading interventions in English for ELLs is limited, a few studies have been conducted.

Gunn, Biglan, Smolkowski, and Ary (2000) randomly assigned early readers at risk for reading problems who were ELLs to treatment and control groups for an English reading intervention. Students in the treatment condition were provided a direct-instruction approach to reading for 25–30 minutes daily for 5 months to 2 years (precise number of intervention sessions provided was not specified). Most students made significant gains on word attack. Students who participated for the entire 2 years made significant gains on letter-word identification, word attack, reading vocabulary, and passage comprehension, but showed no treatment effects for fluency. Data were provided separately for a subgroup of 19 students who were non-English speaking and participated in the treatment. These students made significantly greater gains than controls on words read per minute. No statistically significant gains on other reading outcomes were reported for non-English-speaking students—though other measures favored the treatment group—and power was low to detect differences.

In another study, 26 second graders at risk for reading problems who were ELLs

participated in 58 English reading intervention sessions (35 minutes each) in groups of one to three (Linan-Thompson et al., 2003). Students made significant gains from pre- to posttest on word attack, passage comprehension, phoneme segmentation fluency, and oral reading fluency. The absence of a comparison group makes findings more difficult to interpret; however, only three students made less than 6 months' growth during the 3-month intervention. Oral language proficiency in Spanish or English did not predict the performance of the three students who made the least gains. The ELLs who participated in the intervention gained more than two words per week in fluency (on average), whereas on-level English monolingual readers gained slightly above one word per week in fluency (Hasbrouck & Tindal, 1992).

Denton, Anthony, Parker, and Hasbrouck (2004) conducted two types of intervention programs in English for students in grades 2–5 in a bilingual (Spanish/English) program. Students who scored below a first-grade reading level were randomly assigned to a systematic, explicit reading intervention using decodable text or to an untutored control group. Undergraduate students trained by the researchers provided about 22 sessions of tutoring for both sets of students assigned to the treatment condition. Students who were tutored in reading using the intervention significantly outperformed controls on word identification.

These studies (Denton et al., 2004; Gunn et al., 2000; Linan-Thompson et al., 2003) have in common the implementation of reading interventions in English with ELLs, use of interventions that teach phonics and word-level decoding in the context of active engagement with text and comprehension instruction, and significant outcomes in English even for students whose English language skills were developing while they were being taught to read in English. The Denton et al. (2004) study provided initial support for gains at the word level but less support for fluency and comprehension. Though this study yielded gains in overall growth in fluency and comprehension, there was no control group. The Gunn et al. (2000) study resulted in growth in oral reading fluency for ELLs in the intervention group. The studies by Denton et al. and Linan-Thompson et al. provided interventions that were relatively brief (22 sessions and 58 sessions, respectively); the total amount of intervention provided in the Gunn and colleagues study was difficult to discern. These studies did not include instruction in English language development. Also, with the exception of the Linan-Thompson et al. study, it appears that the instructors did not make adjustments to accommodate ELLs.

Design, Development, and Theoretical Background of the Intervention for this Study

English language learners, like monolingual English speakers, learn to read through phonological recoding and spelling-sound patterns (Lopez, 2004; Signorini, 1997; Treiman, 1984). One guiding premise for the design of the intervention was that less skilled students who are learning to read in alphabetic languages have difficulties due to not having mastered the alphabetic principle (Paulesu et al., 2001). Therefore, we designed the intervention to teach the sounds in English and how they relate to letters. Decodable text was used throughout instruction. Simultaneously, instruction addressed learning to read as "sight words" words that were less phonetically regular in English (Ehri & Wilce, 1983; Goswami, 1993; Rayner, Foorman, Perfetti, Pesetsky, & Seidenberg, 2001). Sight words or high-frequency words were introduced prior to students reading them in their stories. We recognized that students could learn to master the code in English with intensive intervention but might not make adequate progress in comprehension unless we aligned the intervention with current research on developing vocabulary and com-

prehension (Beck, McKeown, & Kucan, 2002; Fitzgerald, 1995a; Gersten & Baker, 2000; Snow et al., 1998; Ulanoff & Pucci, 1999).

The intervention reflected the six instructional practices in reading that are effective for beginning readers who are ELLs (Gersten & Geva, 2003): (*a*) explicit teaching, (*b*) promotion of English language learning, (*c*) phonemic awareness and decoding, (*d*) vocabulary development, (*e*) interactive teaching that maximizes student engagement, and (*f*) instruction that produces opportunities for accurate responses with feedback for struggling learners. The foundational skills of phonemic awareness and phonics are more critical in the beginning stages of reading and less important as students become readers of connected text. Improving vocabulary and word knowledge is an important part of reading and all content learning throughout the instruction. Reading comprehension was initially taught through listening comprehension and later through reading text.

Our instructional design provided systematic and explicit instruction for teaching phonemic awareness, phonemic decoding skills, fluency in word recognition and text processing, construction of meaning, vocabulary, spelling, and writing (see Foorman & Torgesen, 2001; National Reading Panel, 2000; Pressley, 1998; Rayner et al., 2001; Snow et al., 1998). We designed explicit instruction by scaffolding instruction so that students are not required to infer new knowledge but rather the teacher conveys new knowledge. Intensive instruction is reflected in activities that require high student engagement in learning critical content and in the relatively small (3–5) ratio of teachers to students.

We examined the effectiveness of a 7-month daily intervention (50 minutes per day) for ELLs (Spanish/English) whose primary reading instruction was in Spanish (Vaughn et al., 2006). We designed and implemented an intervention aimed to improve students' reading and language skills in Spanish. First graders at high risk for reading problems from the same schools and classrooms were randomly assigned to our treatment or a control condition. Treatment students outperformed control students on phonological awareness, word attack, fluency, comprehension, and oral language outcome measures. Students in the treatment condition who were very low in all reading-related skills at pretest scored above average (word attack, comprehension), average (dictation), or near average (fluency, language composite) at the end of the intervention.

The intervention described in this article was also designed for first-grade ELLs (Spanish/English) at risk for reading problems. Participants in this study were at-risk first graders who were receiving initial reading instruction in English. This approach for ELLs is being used increasingly across the United States and has been mandated in several states (e.g., California, Massachusetts, and Arizona) and school districts. The intervention was matched to the language of students' core reading instruction (English). We were interested in the effectiveness of an English intervention for at-risk students randomly assigned by school and class to treatment or control conditions. We also included an English oracy component (story reading with retelling and vocabulary development) with the reading intervention to improve English comprehension and vocabulary. Specifically, we explored two questions. First, we wanted to know if struggling readers who were native Spanish speakers but learning to read in English made greater growth in reading and language when participating in intense, small-group reading interventions in addition to their core reading instruction in English. Second, we examined how well the reading skills the students developed in English influenced skills in their native language (i.e., Spanish).

Method

Participants

School sites. The study was conducted as part of a much larger research program

addressing bi-literacy and oracy development in bilingual students (Spanish/English). Because we were interested in the effectiveness of the intervention within contexts that supported literacy acquisition, we focused on effective schools (determined by performance of students in the school on state-wide reading assessments). We reasoned that this would increase the likelihood that students identified represented real reading difficulties and not just lack of adequate reading instruction.

From two districts in Texas that serve large numbers of bilingual students (border district, large urban district) we selected four schools that were considered effective for bilingual students using a priori selection criteria: (*a*) schools were providing English intervention for reading to at least two classes of first-grade students who were ELLs, (*b*) at least 60% of the population of the entire school was Latino, and (*c*) 80% or more of third-grade students passed the state reading achievement tests. The average percentage of Hispanics across these four schools was 98. The population of Spanish-speaking ELLs in kindergarten and first grade ranged from 48% to 99%. All schools participated in the free or reduced-price lunch program, and the proportion of students who qualified ranged from 85% to 100%.

Students. We screened students from 14 first-grade classrooms across the four schools at the beginning of first grade (4 weeks after school began). Two tests in both Spanish and English were used for the screen, for a total of four measures: (*a*) the letter-word identification (LWID) subtest from the Woodcock Language Proficiency Battery (English and Spanish, described below), and (*b*) the first five words from a word-reading list (English and Spanish) used to assess initial word-reading ability. We developed this word list by first generating words from kindergarten to grade 3 instructional cumulative vocabulary materials. The list was then matched to the LEXESP database of printed-word frequencies in Spanish (Sebastián, Cuetos, Martí, & Carreiras, 2000), which is similar to the Zeno, Irens, Millard, and Duvvuri (1995) database of printed word frequencies in English. Forty words were selected from the kindergarten to grade 3 instructional corpus with varying probability depending on grade and printed-word frequency. Specifically, we selected words with low probability if they came from the kindergarten corpus, whereas the probability of selection for nonkindergarten words varied by the log of a word's printed frequency in grade 1 to 3 texts (higher for low-frequency words, lower for high-frequency words).

The final list consisted of 40 words representing a diversity of linguistic features, ordered by difficulty to span kindergarten to third grade. This measure had high reliability in grades kindergarten through 3 (internal consistency was over .90 in each grade in the parent project sample of roughly 4,003 students). The first five words are the easiest and consist of two- to four-letter words (in English the words were: in, as, it, man, dog; in Spanish the words were: el, las, un, por, alto). Criteria for selecting students as eligible for the intervention were scoring below the twenty-fifth percentile for the first grade on the LWID subtest in both Spanish and English, and inablilty to read more than one word from the list of simple words. The rationale for assessing students in both Spanish and English (as opposed to English only) was that, given differences in children's exposure to English print materials, poor performance in English but not Spanish could be due solely to a lack of exposure to English reading rather than to a student possessing characteristics placing him or her at risk of developing language and learning difficulties.

Two hundred sixteen students were administered both the Spanish and English screen at the four target schools (although two students refused both English screening tests, one student refused one English subtest, and one student refused one Spanish subtest). Of the remaining 212 students,

109 (51%) met the Spanish intervention inclusion criteria, 65 (31%) met the English intervention inclusion criteria, and 56 students (26%) met both criteria; thus, nine students met the English cutoff but not the Spanish cutoff, and these students were not eligible for the intervention. Of the 56 students meeting criteria for both English and Spanish, eight were not available for randomization because five had left the school, and three other students declined to participate. The remaining 48 students were randomly assigned within schools to either supplemental intervention or contrast conditions. The composition of the randomized groupings changed for five intervention students where scheduling was not possible, and each of these was replaced by an alternate within that school, resulting in 10 cases of failed assignment. It is important to note that these assignment failures occurred prior to the onset of the intervention and resulted from the inability to accommodate students' schedules given the small-group nature of the intervention.

As discussed later in the results section (see Table 1), there were no statistically significant differences on the pretests between intervention and contrast children on any measure, in either language. The mean effect size difference between intervention and contrast groups was $d = .21$ (range .02–.42), and all confidence intervals comfortably included zero; nine of 30 differences favored the contrast group. Across groups, decoding skills were low but stronger in English relative to Spanish; oral language performance was also low but more similar across languages. This general pattern of pretest findings (in terms of group differences as well as performance levels) pertained with or without the failed-assignment students as well as when comparing treatment and contrast groups based on their originally randomized composition. For posttest results, although an intent-to-treat analysis was not practical given the small sample size, we did analyze primary results both with and without these students with no substantive differences noted; therefore, subsequent results included all children who did and did not receive the intervention.

The study was initiated with 24 intervention students and 24 contrast students and, due to ordinary attrition (in each case, students' families moved away from their home school), the study ended with 22 intervention and 19 contrast students (8% attrition for intervention and 21% attrition for contrast); in addition, data were not obtainable for one contrast student at either testing time. The mean age of the 47 students with pretest data was 6.59 years ($SD = 0.54$). All students were Hispanic, and female students comprised 50% of the sample ($n = 23$).

Classroom teachers. Thirteen teachers provided core reading instruction in English to participating students. Teachers averaged 11.8 ($SD = 8.3$) years teaching overall and 7.3 ($SD = 5.7$) years teaching first grade. Nine (69%) were Hispanic, 10 (77%) were female, and all had at least a bachelor's degree. Overall, eight (62%) had credentials as bilingual teachers, and four (31%) were certified to teach English as a second language. Randomization was carried out within schools, but did not explicitly stratify on classroom teacher or other factors (e.g., pretest). That is, assignment was carried out at the student level within schools without stratification.

Measures

A comprehensive battery of language/literacy-related measures in English and Spanish was administered prior to the onset of intervention (October) and following its completion (May).

Letter naming and letter-sound identification. Students were asked to identify and provide at least one sound for each of the 26 letters of the English alphabet and each of the 30 letters of the Spanish alphabet. Dependent measures were the raw score totals for each measure.

Comprehensive Test of Phonological Processing (CTOPP; Wagner, Torgesen, &

Rashotte, 1999). The CTOPP has nine subtests measuring phonological awareness (PA), rapid naming (RN), and phonological memory (PM). The normative base consisted of 1,656 individuals from ages 5 to 24, similar to the 1997 U.S. Census statistics. Coefficient alphas for all three composites in the entire normative sample ranged from .83 to .95, and .83 to .92 in the age range of this sample; test-retest estimates in a sample of 32 children ages 5 to 7 ranged from .70 to .92 for the three composites. In addition, content, concurrent, predictive, and construct validity data are provided in the CTOPP manual (Wagner et al., 1999).

We used seven subtests, including elision, blending words, blending nonwords, segmenting words, sound matching (first sound and last sound), nonword repetition, and rapid letter naming (Form A or B). Elision consists of 20 items and requires a student to listen to an audiotape, repeat a real word, and then repeat the word with a specified phoneme deleted, which may appear in the initial, middle, or rime portion of the word; in each case, the result is a real word. Blending words has 20 items requiring a student to blend several phonemes heard from an audiotape into a real word; items increase in difficulty by an increase in the number of phonemes to be blended from two to eight. Blending nonwords has 18 items and is identical to blending words, except that the final result is a nonword. Segmenting words has 20 items and requires a student to listen to an audiotape for a word and then to identify the individual phonemes that comprise the word; items increase in difficulty by an increase in the number of phonemes within a word, which ranges from two to nine. Sound matching contains two parts, first sound and last sound, each consisting of 10 items, and requires a student to look at a picture book that displays three pictures and then to identify the word that starts (or ends) with the same sound as an exemplar. Nonword repetition has 18 items and requires the student to repeat a nonword that increases in length and complexity, heard from an audiocassette. Finally, rapid letter naming requires a student to read from left to right and line to line a series of six letters arranged in a 9 by 4 format.

Although age-based standard scores are available for the CTOPP, we used raw scores in analyses to compare performance with a Spanish-language version of this instrument (see below). For rapid letter naming, scores were converted to the number of letters identified per second. We administered the blending-words measure first to reduce frustration and decrease testing time. If a child received a raw score of zero on this measure, he or she was administered sound matching but not the other CTOPP measures; in this case, full credit was given for the sound-matching subtest. If students received a raw score of at least 1 on blending words, they were not administered sound matching but were given the other subtests. In these cases, scores of zero were imputed based on the low likelihood of a nonzero total given the greater difficulty of the omitted subtest. In addition, for any subtest, if a child was unable to complete any of the practice items, the test was not given and a score of zero was imputed. These decisions were derived from earlier research with the predecessor to the CTOPP and from empirical modeling of performance on this test using item-response methods (Schatschneider, Francis, Foorman, Fletcher, & Mehta, 1999) as well as work in the parent project examining the properties of this assessment in a larger sample (n = 1,600) of bilingual students.

A phonological awareness (PA) composite was created from the subtest scores of the sound matching, blending words, blending nonwords, segmenting words, and elision subtests of the CTOPP. Where the sound matching subtest was not administered because students met performance criteria on the blending words subtest, a perfect score was imputed for purposes of calculating this composite. (Analyses were also performed where a composite was

computed without the sound matching subtest, and results were highly similar.)

Test of Phonological Processes—Spanish (TOPP-S). The TOPP-S (Branum-Martin et al., 2006) was developed to align with the English CTOPP in terms of skills assessed and the linguistic complexity of the items within each subtest while still being appropriate for the Spanish language. Each subtest consists of comparable numbers of items as those in the CTOPP. With the exception of sound matching, all subtests were built entirely of production-based items, and items were targeted to match CTOPP items in linguistic complexity (e.g., number of phonemes, area of manipulation) but relied on phonemes and syllables appropriate for the Spanish language. Reliability estimates for the TOPP-S were determined on a sample of approximately 1,500 students, and the coefficient alphas were very high, ranging from .93 to .97. We used raw scores comparable to those calculated for the CTOPP for data analyses; the same branching rules for the CTOPP were also used for the TOPP-S. In addition, a phonological awareness composite in Spanish was created using the analogous subtests and the same rules as the English PA composite.

Woodcock Language Proficiency Battery—Revised: English and Spanish Forms (WLPB-R; Woodcock, 1991; Woodcock & Munoz-Sandoval, 1993). The WLPB-R is a standardized instrument whose normative sample was concordant with 1980 U.S. Census statistics, consisted of 6,359 subjects (3,245 in K to 12), and was the same as that of the Woodcock-Johnson Psychoeducational Battery—Revised (Woodcock & Johnson, 1989). Median coefficient alphas range from .81 to .92 across all age ranges (and from .77 to .96 at ages 6 to 9) for the subtests used; test-retest measures for selected subtests in a sample of 504 ranged from .75 to .95. In addition, content, concurrent, and construct validity data are available in the WLPB-R manual (Woodcock, 1991). The Spanish Form was derived from 3,911 native Spanish-speaking individuals from 22 countries, including 1,325 from the United States and 1,512 from Mexico who were close to monolingual Spanish speakers; median coefficient alphas range from .84 to .92 across all age ranges and from .68 to .95 at ages 6 and 9 (Woodcock & Munoz-Sandoval, 1993). The test development, scaling, and norming process for the assessment is described in detail in the WLPB-R manuals (Woodcock, 1991; Woodcock & Munoz-Sandoval, 1993) along with content, concurrent, and construct validity data. To obtain scaled scores on either the Spanish or English language assessment, the raw score for a given subtest in Spanish or English is converted to a score on the W scale (a Rasch ability score, which is a variant of an item-response theory scale), and then the difference between this W score and a reference W score (which is based on the child's age but is invariant across languages) is used to determine the student's age- and language-specific standard score for that measure (Woodcock, 1991; Woodcock & Munoz-Sandoval, 1993). This scaling process allows scores on the English and Spanish language assessments to be compared in the sense that it places the Spanish language norms on the same scale as the English language norms.

Subtests used for this study were letter-word identification (at screening only), word attack, passage comprehension, listening comprehension, picture vocabulary, verbal analogies, and memory for sentences (at pretest only). Letter-word identification requires the student to match a rebus to a picture of an object (beginning items), then to read aloud individual letters, and then to read aloud words that increase in length and complexity. Word attack requires the student to read aloud nonsense or unfamiliar words that are linguistically logical. Passage comprehension first requires students to point to a picture represented by a phrase in a multiple-choice format and then to read a sentence or short passage and provide a missing word that is appropriate for the context of the passage. Lis-

tening comprehension is similar to passage comprehension in the oral domain and asks the student to listen to a passage and supply the missing word at the end using an oral cloze procedure. Picture vocabulary requires the student to name familiar and unfamiliar pictured objects and is primarily an expressive semantic task. Verbal analogies asks a student to provide verbal answers to questions about logical relationships that increase in difficulty. Finally, memory for sentences requires a student to repeat phrases or sentences that increase in length.

Dependent measures were age-based standard scores only, although we analyzed raw scores with similar results. In addition to the individual subtests described above, an oral language composite was also calculated based on the average W score of the measures available at each testing point. At pretest, these measures included memory for sentences, listening comprehension, picture vocabulary, and verbal analogies; at posttest, these measures included listening comprehension, picture vocabulary, and verbal analogies.

Dynamic Indicators of Basic Early Literacy Skills (DIBELS; Good & Kaminski, 2002)/***Indicadores Dinámicos del Exito en la Lectura*** (Good, Bank, & Watson, 2003). This is a measure of reading fluency requiring the student to orally read a passage geared to the student's grade level. Children were given a maximum of 3 seconds per word and a maximum of 60 seconds to read the entire passage. At pretest and posttest, the first-grade, beginning-of-year passage was administered in both Spanish and English. The dependent measures were the number of words correctly read within the time limit. In addition, at posttest, we also administered the first-grade end-of-year passage in both Spanish and English. The dependent measures were the number of words correctly read, minus the number of words read incorrectly, during the 1 minute allowed for reading.

Intervention

Four bilingual teachers provided the intervention in English to treatment participants in six small groups of three to five students for 50 minutes a day, 5 days a week, from October through May. The intervention was provided in addition to the students' core reading lessons (also in English) and did not conflict with their daily scheduled reading time. The mean amount of time documented for delivery of the supplemental instruction across the 22 students who completed the intervention ranged from 76 to 115 hours ($M = 96.55$ hours, $SD = 11.02$).

The four intervention teachers received 12 hours of professional development prior to teaching and an additional 6 hours after 6 weeks of implementation. Teachers also participated in frequent 1–2-hour staff development sessions and on site coaching during which they (*a*) were provided feedback about their instruction based on observations and videotaped lessons, (*b*) discussed any questions or challenges regarding implementation of the intervention, and (*c*) collaborated in planning and instruction by using case studies from the participants to plan for accelerating the growth of students. These sessions occurred weekly for the first 2 months of implementation and less frequently as intervention teachers improved in confidence and performance.

Reading curriculum. In this research we used a beginning reading curriculum (Mathes, Torgee, Wahl, Menchetti, & Grek, 2004) based on the direct instruction approach (Carnine, Silbert, & Kame'enui, 1997) and designed and validated in previous intervention research with monolingual English struggling readers. Our team modified the English curriculum by interspersing language-support activities appropriate for ELLs throughout the instructional sequence. This Spanish intervention was designed to be a comprehensive, integrated supplemental intervention curriculum that details for teachers how to deliver explicit phonemic awareness and phonics instruc-

tion, teach students to apply alphabetic knowledge to word reading, and engage children in reading for meaning. Instruction was delivered to small homogeneous groups of three to five struggling readers. Each lesson (120 total) required approximately 40 minutes to complete.

Building on the instructional design principles of the model of direct instruction (Carnine et al., 1997), we analyzed the tasks associated with fluent, meaningful reading and ordered elements into a cumulatively building and carefully integrated scope and sequence. For example, elements such as /b/ and /d/, or /r/and/er/ were presented several weeks apart. Similarly, we presented elements that promoted faster movement into word building and text reading earlier (e.g., higher-frequency letter-sound correspondences, high-frequency sight words, closed-syllable words), with elements with lower utility being presented later.

Lesson format: From this scope and sequence, we developed daily lesson plans, each made up of 6 to 10 short activities representing five content strands: phonemic awareness, letter knowledge, word recognition, connected text fluency, and comprehension strategies. Instruction was delivered at a fast pace with numerous interchanges between the teacher and students. In a typical activity, the teacher asked all students to respond to letters, words, or text in unison, followed by "individual turns" where each child was able to demonstrate his or her ownership of the content. Teachers modeled new content, providing guided practice for students and implementing independent practice for every activity. Teachers consistently monitored students' responses, providing feedback for correct answers and scaffolding learning when responses were incorrect.

Instructional content: In a typical lesson, students played word games designed to promote phonemic awareness, practiced letter-sound correspondences for previously taught letters or letter combinations, practiced writing these letters, and learned the sound of a new letter or letter combination. Students also practiced sounding out and reading words composed of previously taught letter-sound correspondences and various syllable types, spelled words from dictation based on their sound-symbol correspondences, practiced automatic recognition of words that do not conform to alphabetic rules, read and reread decodable connected text, and applied comprehension strategies to this text.

Phonemic awareness: The phonemic awareness strand included two types of activities: phoneme discrimination, and phoneme segmentation and blending. Early activities required children to isolate initial sounds in words or to tell if a word started with a particular sound. Later these activities moved to isolating final and medial sounds. Sound-discrimination activities were also used to ensure that children were sensitive to the differences in vowel sounds. Likewise, children were taught how to segment one-syllable words into individual phonemes as well as to reconstitute words from individually spoken phonemes. When students were able to segment and blend words containing consonant blends, this strand ended.

Letter knowledge: Letter-sound correspondences were introduced from the first day of instruction and continued through all 120 lessons, with a new letter-sound or letter-combination-sound correspondence being introduced every 2 to 3 days. Prior to presenting the letters representing a phoneme, that phoneme was manipulated orally during segmenting and blending activities. The primary objective of the letter-knowledge strand was to develop automatic recognition between a letter symbol and the most common sound it represented. Also, students were asked to both say the phoneme represented by each letter or letter unit and to write letters as the teacher dictated phonemes. When students knew more than one way to write a phoneme, they were asked to write it both ways. For ex-

ample, in later lessons children were expected to write /s/ as s, ce, and ci.

Word recognition: The word recognition strand taught both phonetically regular and irregular/high-frequency words. In terms of decoding phonetically regular words, students were initially taught the sounding-out strategy with CVC words (e.g., man, fit, cup) and later with more complex words. As students moved toward decoding unknown words quickly and efficiently, they were also learning to read words representing six syllable types and learning to read multisyllable words. Initially students applied the sounding-out strategy to each syllable, read each syllable "fast," then read the whole word. The sounding-out step was quickly removed so that children read each syllable part, then read the whole word. By the end of the program, children were reading two- and three-syllable words using all six syllable types.

Teachers presented high-frequency words that were irregular as "tricky words that don't sound-out right." Even so, children were asked to sound them out, followed by analysis of the parts that "worked right" and the parts that had to be remembered.

Connected text fluency: Students practiced application of word recognition strategies through the daily reading of decodable text (all phonetic elements and all irregular sight words appearing in the text had been taught previously). To facilitate fluency, children were asked to read stories two to three times, with a goal of improving rate and accuracy. Typically children read a story in unison on the first reading. On the second reading, children usually read one to two pages of a story individually. The third reading was typically read in pairs, with the teacher pairing with one child and timing that child's reading rate.

Comprehension: The goal of the intervention was for students to read connected text rapidly and with comprehension. Thus, prior to reading a story each day, teachers engaged in "browsing the story" and teaching students to make predictions. Teachers and students then established a purpose for reading. With expository text, teachers activated prior knowledge by asking students to tell what they already knew about the topic. After reading the story, students then engaged in a number of reading comprehension activities including story retell, sequencing, and summarizing. As students advanced they were taught story grammar elements, or, with expository text, how to identify new information learned.

Language support modifications: To ensure the students would understand and benefit from the *Proactive Reading* curriculum, we made language modifications to make concepts comprehensible. Thus, we created language support activities that were interspersed throughout each proactive lesson. The number of language support activities for each lesson varied from three to eight and included instructional scripts with pictures. Additionally, instructional practices identified as effective when working with students who are ELLs were incorporated in the lessons, including use of visuals, gestures, and facial expressions in teaching vocabulary and clarifying meaning of content; provision of explicit instruction in English language use; and opportunities to give elaborated responses.

To ensure that students understood the tasks they were asked to perform, we defined words in the directions for performing a task they might not know (e.g., *trace*, *copy*) prior to beginning the task. Before each lesson in which teachers used word lists to complete tasks or in which students would read connected text, words from the list or text that students may not have known were also defined. Students were told that the words would be used in the task or would appear in the story. If students had previously learned a specific meaning of a word, teachers reminded students of the meaning they had learned, told students they would learn a new meaning, and provided a sentence as a prompt. The same was true if students had learned the target word as a particular part of speech.

Oracy and vocabulary development: Because participating students were bilingual, at risk for reading problems, and had low language proficiency scores in both English and Spanish (see Table 1), we emphasized the development of oracy and vocabulary in English (the language of reading instruction). Although oracy and vocabulary development were integrated into the reading instruction, we assigned 10 minutes each day to the development of vocabulary, listening comprehension, and language development. All instructors used the same largely expository books ($n = 25$) in English centered on eight information themes (e.g., pets, bugs). The first theme was the only narrative theme, and it addressed families. Each theme was addressed through reading three or four books on the topic. We selected two to three key vocabulary words each day that corresponded with the text that the teachers were reading, and teachers taught the meaning of the words and used them in sentences with students prior to reading the passage from the book to students. Teachers read passages to students each day and then asked questions about the vocabulary and key ideas. Teachers used probes to guide students in story retelling, providing opportunities for each student to participate. Students discussed the story with the teacher using complete sentences and new vocabulary terms.

Intervention Instructors and Intervention Validity Checks

The four intervention instructors were bilingual (Spanish/English), had at least an undergraduate degree, were hired by the research team, provided the intervention outside of core reading instruction, and were well prepared to do so (see Intervention subsection).

Across the year, two observers, in consultation with the primary author of the intervention, worked closely to obtain interrater reliability using videotapes of bilingual intervention teachers implementing the English intervention curriculum. Upon obtaining interrater reliability of 95%, both observers conducted intervention validity checks during the beginning, middle, and end of the year so that each instructor was observed three times for fidelity of implementation. Interrater reliability was reestablished prior to each intervention validity check.

We developed the intervention validity instrument to collect both quantitative and qualitative data on the following teacher behaviors at each observation time: (*a*) pacing instruction, (*b*) providing independent practice, (*c*) presenting the lesson appropriately, (*d*) providing error correction, (*e*) providing appropriate scaffolding, (*f*) teaching concepts to mastery, (*g*) maintaining student attentiveness, and (*h*) eliciting student responses (Grek, Mathes, & Torgesen, 2003). Using guidelines, observers assigned a numerical rating of 1–3 to each of the eight areas listed above for every activity observed: 1 = poor: the instructional behavior greatly deviated from specified guidelines; 2 = average: the instructional behavior met most but not all guidelines specified; or 3 = excellent: the instructional behavior met all guideline specifications. Observers also wrote field notes to provide further details on each instructional behavior described above. Across numerous activities and the three observations, the average rating scale score (maximum possible = 3) for teachers providing the intervention ranged from 2.38 to 2.65 with an overall average of 2.58 ($SD = 0.13$), corresponding to a ranking between average and excellent.

Using a list of nine questions that addressed general teacher preparedness to teach the intervention, intervention teachers were also rated as "yes" (the item was present or observed) or "no" (the item was not present or observed). The nine questions addressed (*a*) materials ready, (*b*) materials visible to students, (*c*) students seated appropriately, (*d*) instructor's enthusiasm/warmth, (*e*) ongoing monitoring of student performance, (*f*) checking of practice items for correctness and providing

feedback, (*g*) redirection of off-task behavior, (*h*) communication of clear expectations and learning goals for activities, and (*i*) participation of each student during the story retell. Across independent observations, instructors received an average of 96% "yes" responses.

Core Reading Program and Classroom Observations

Core reading program. The core curriculum used in the schools for reading during the classroom reading instruction for the urban border city was Language Enrichment (Carreker, 1999) and in the large urban city was McGraw Hill Reading (McGraw Hill, 2001).

Observation measure. Classroom literacy instruction was observed using an observation scheme developed by Foorman and colleagues (Foorman, Goldenberg, Carlson, Saunders, & Pollard-Durodola, 2004; Foorman & Schatschneider, 2003) to record time allocated to various activities during reading instruction. Using the framework, observers make on-the-minute observations of the teacher and students during reading/language arts and English language development. Observations were conducted three times across the year (beginning, middle, and end). Mean reliabilities were 80% or higher. Trained bilingual researchers not working on this intervention study and unfamiliar with which students were assigned to treatment and control conditions collected all data.

The time per day classroom teachers reported that they taught reading/language arts was 144 minutes (range 120 to 180 minutes), and this was equivalent to the average time observed in these classrooms ($M = 146$, $SD = 29$). Approximately 92% of the time observed consisted of instructional time. During the instructional time observed, there was a variable distribution of instruction across the following reading-related categories (in descending order): word work ($M = 23.0\%$, $SD = 10.9\%$), oral language ($M = 16.7\%$, $SD = 8.3\%$), reading ($M = 12.8\%$, $SD = 6.5\%$), writing/spelling ($M = 9.4\%$, $SD = 3.6\%$), and reading comprehension ($M = 6.8\%$, $SD = 3.9\%$). The remaining instructional time was spent giving directions ($M = 14.6\%$, $SD = 3.5\%$), providing feedback ($M = 4.3\%$, $SD = 3.7\%$), and in nonreading instruction ($M = 12.5\%$, $SD = 4.1\%$).

Although all core reading instruction was provided in English, a minor proportion provided in Spanish was also observed. On average teachers used English 78.1% of the time ($SD = 10.4$), Spanish 4.4% ($SD = 7.1\%$), and a mix of English and Spanish 5.7% of the time ($SD = 7.1\%$); some instructional time did not involve oral language or was not codable.

Reading intervention for contrast students. We individually interviewed classroom teachers to determine the extent to which contrast students were provided reading interventions (instruction in addition to the core) by the schools. A member of our research team interviewed each intervention student's classroom teacher three times over the school year. A standardized form with questions was completed for each student to determine the type and amount of additional instruction, if any. Fourteen contrast students received one or more types of reading intervention in addition to their core reading instruction. The total time supplemental reading intervention was provided to these 14 students ranged from 24 to 259 hours ($M = 63.7$ hours, $SD = 61.1$ hours; median = 47.3 hours); the overall median for contrast students was 23.8 hours. Of the 24 intervention students, 12 received supplemental reading intervention (outside of the intervention study), and the total amount of time spent in these activities ranged from 23 to 62 hours ($M = 35.7$ hours, $SD = 13.7$ hours; median = 32.5 hours); the overall median for all 24 students was 11.4 hours.

Analyses

We completed several sets of analyses. First, preliminary analyses were conducted

for all subtest scores to examine performance distributions. We also compared intervention and contrast groups on all dependent measures prior to the onset of intervention. The next set of analyses examined posttest performance as a function of group (intervention or contrast), controlling for pretest performance as well as age or other covariates when necessary (see below). Standardized effect sizes (Cohen's *d*; Cohen, 1988) were also computed, using differences in mean performance divided by the pooled original standard deviation at posttest for each measure; these values were then adjusted for sample overestimation bias (Hedges & Olkin, 1985), and we calculated confidence intervals (95% limits) based on the standard error of the corrected *d* values. Because the sample sizes of the intervention and contrast groups were not optimal, we also conducted power analyses for the posttest results using the Power and Precision v. 2.1 software (Borenstein, Rothstein, Cohen, Schoenfeld, & Berlin, 2001). These analyses were conducted where nonsignificant statistical differences arose between groups. Specifically, power for detecting current group differences (at the .05 level) was calculated based on the present sample parameters (*N*, mean, *SD*, effect size); the sample size necessary for 80% power was also determined. Power analyses were not conducted for results where statistically significant differences occurred because adequate power was assumed, because a smaller sample size biases against finding statistical differences.

Results

Preliminary Analyses

Examination of preintervention score distributions through box and stem-and-leaf plots, as well as other univariate statistics, indicated adequate variability on most subtests assessing phonemic awareness, language, and word reading skills in English and Spanish; floors were apparent on the experimental word reading lists in both languages, which is not surprising given the selection criteria for the study. Floors were also noted on measures of reading fluency (rapid letter naming and DIBELS) in both languages; in addition, most students knew the letter names in English, but scores were skewed to the low end for Spanish letter naming, as well as for WLPB-R word attack performances in Spanish. Average age did not differ between the intervention and contrast groups, $F(1, 45) < 1$. Most students were age appropriate for first grade in Texas (6 years old), although one intervention student was significantly older. Group-difference analyses (at pretest and posttest) were conducted with and without covariation for age, but no substantive differences in results were observed. Results reported in subsequent sections therefore are uncorrected for age.

Preintervention Performance (English and Spanish)

Pretest performance means for intervention and contrast groups are presented in Table 1. As expected, given that students were assigned to the groups randomly, there were no significant group mean differences in performance on either of the skills (WLPB-R letter word identification and experimental word reading list) used in the intervention screen, in either English or Spanish. Furthermore, as noted earlier, mean comparison of skill performance on the larger battery administered prior to the onset of treatment indicated that students in the intervention and contrast groups performed at comparable levels on all English and Spanish skills assessed, with no significant differences between groups on any measure. In addition, effect-size differences between groups were also small. Reading and language performances were approximately 1 to 3 standard deviations below normative levels for both groups, with performance nearing the average range only for English word attack scores (for both groups).

In general, letter and word naming and reading (e.g., letter naming, rapid letter

TABLE 1. Pretest Perfomance on Language and Reading Measures in Spanish and English

Measure/Group	English			Spanish		
	n	*M*	*SD*	*n*	*M*	*SD*
Letter naming:						
Letter name identification (raw score):						
Intervention	24	12.58	7.66	24	18.25	8.71
Contrast	22	11.27	9.23	23	15.87	8.87
Rapid letter naming (CTOPP/TOPPS-S; letters per second):						
Intervention	23	.06	.21	24	.42	.42
Contrast	23	.09	.21	23	.38	.40
Phonological processing:						
Letter sound identification (raw score):						
Intervention	24	14.83	6.07	24	16.08	7.25
Contrast	22	12.68	7.75	23	14.74	7.63
Nonword repetition (CTOPP/TOPPS raw score):						
Intervention	23	8.65	3.96	24	4.67	3.29
Contrast	23	7.13	4.43	23	3.74	3.56
Phonemic awareness composite (% correct):[a]						
Intervention	24	.31	.12	24	.31	.14
Contrast	23	.29	.14	23	.26	.17
Language (Woodcock, standard score; $M = 100$, $SD = 15$):						
Listening comprehension:						
Intervention	22	76.05	15.6	23	49.22	22.8
Contrast	23	65.83	29.6	22	50.50	20.1
Picture vocabulary:						
Intervention	22	58.41	23.6	23	62.17	22.7
Contrast	22	60.45	29.1	23	56.09	25.2
Verbal analogies:						
Intervention	22	85.23	15.2	23	80.52	15.9
Contrast	23	80.52	19.6	21	77.10	14.6
Memory for sentences:						
Intervention	23	67.96	16.2	23	71.17	13.3
Contrast	23	69.83	14.9	23	66.39	16.2
Oral language composite:						
Intervention	22	68.82	17.1	23	60.87	19.4
Contrast	23	63.52	26.0	23	54.78	22.0
Reading:						
Woodcock letter word identification (raw score):						
Intervention	24	7.17	3.48	24	9.83	3.42
Contrast	24	7.25	3.00	24	9.71	3.42
Word attack (standard score; $M = 100$, $SD = 15$):[b]						
Intervention	20	61.25	17.7	18	89.00	12.4
Contrast	16	66.94	13.2	17	87.88	8.72
Passage comprehension (standard score; $M = 100$, $SD = 15$):[b]						
Intervention	17	75.41	8.60	20	83.70	15.5
Contrast	14	78.71	9.33	16	82.75	8.15
Experimental word list (raw score):						
Intervention	24	.00	.00	24	.25	.44
Contrast	24	.04	.20	24	.21	.41
DIBELS BOY (raw score):[b]						
Intervention	16	1.19	2.10	16	1.50	2.53
Contrast	17	.82	.95	14	1.79	2.08

NOTE.—There were no pretest differences (all $p > .65$) between groups on any measure in either language. DIBELS BOY = Dynamic Indicators of Basic Early Literacy Skills Beginning of Year Story.

[a] Includes scores from the sound matching, blending words, blending nonwords, segmenting words, and elision subtests of the CTOPP/TOPPS.

[b] Measures for which data were not present at pretest (due to inability) but available at posttest.

naming, letter-sound identification, letter-word identification, word attack, passage comprehension) appeared higher for both groups when measured in English relative to Spanish. For example, the average percentile rankings across groups in English for word attack and passage comprehension were at the twenty-third and fourteenth percentiles, respectively (mean standard scores = 88.4 and 83.3), whereas the average percentile rankings across groups in Spanish for word attack and passage comprehension were below the first (mean standard score = 63.8) or the sixth (mean standard score = 76.9) percentiles, respectively. In contrast, language skills (nonword repetition, phonological awareness composite, WLPB-R oral language composite) across languages were generally comparable. Specifically, although students' overall average language ability (oral language composite) was below the first percentile in both languages, there was some variability within the individual subtests, with average percentile rankings across groups and languages at or below the first percentile for memory for sentences and picture vocabulary, and at or below the thirteenth percentile for verbal analogies; for listening comprehension, percentile rankings across groups in Spanish were at approximately the second percentile (mean standard score = 70.8), and for English, well below the first percentile (mean standard score = 49.8).

As can be seen in Table 1, for three measures (WLPB-R word attack and passage comprehension subtests, and DIBELS), the sample size was markedly lower than for the other measures (in both languages, and for both groups) because children refused or were unable to provide data on these measures at the beginning of their first-grade year.

Posttest Performance: English Measures

Results of posttest performance in English are presented in Table 2. Included are means, effect sizes, significance tests, and gain scores for students who had test data at both testing times. Performances are discussed by area.

As earlier noted, a portion of the sample did not have pretest data for English word attack, passage comprehension, and/or DIBELS; however, most of these students provided posttest data. Using pretest performance as the covariate for posttest performance on the same measure would unduly limit the sample to students who had data at both times. Instead, we used the reading-related measures at pretest that were most predictive of pretest performance on the three measures (for those who did have data at pretest) as covariates for posttest performance. For English word attack, these covariates were (English) letter-word identification, phonological awareness composite; for English passage comprehension, the covariate was (English) letter-sound identification. For English DIBELS (beginning-of-year story), the covariate was the (English) CTOPP rapid letter naming subtest. This subtest was also used as the covariate for the DIBELS end-of-year story (because this story was given only at posttest).

Letter naming and letter naming fluency. Intervention and contrast students did not differ in their ability to name English letters ($p > .05$) after adjusting for pretest performance on this measure; both groups were near ceiling at posttest. However, the effect size of the difference did favor intervention students (+0.59), and power was only 0.45; a sample size of 45 per group would have allowed for power to detect a difference ($p = .05$) of 0.80. Performance on the CTOPP rapid letter naming subtest indicated that, after adjusting for pretest performance, intervention students' letter naming fluency was greater than that of the contrast students, $F(1, 37) = 9.68$, $p < .004$; the standardized effect size of the difference between groups was large (+.88).

Phonological processing. The phonological awareness (PA) composite was created from the subtest scores of the sound matching, blending words, blending nonwords,

segmenting words, and elision subtests of the CTOPP. Performance on the PA composite measure indicated that intervention students outperformed contrast students on these measures after adjusting for pretest performance, $F(1, 37) = 18.51$, $p < .0002$; intervention students correctly answered an average of 60% of the items across the subtests administered, relative to 46% for contrast students, and the standardized effect size of the difference between groups was large (+1.24). Results were similar for performance on the letter-sound identification subtest, $F(1, 37) = 12.50$, $p < .001$, effect size = +1.10. On a measure of phonological memory, however, results were different. On the CTOPP nonword repetition subtest, intervention and contrast students did not differ after adjusting for pretest performance, $F(1, 37) = 1.99$, $p > .05$, with a modest effect size favoring intervention students (+.46), and a power of only 0.30; a per-group sample size of 72 would allow for power of 0.80.

Oral language. On the WLPB-R oral language composite, analyses indicated that intervention and contrast students did not differ after adjusting for pretest performance, $F(1, 36) < 1$, $p > .05$, with a modest effect size (+0.43), but with power of only 0.27; a per-group sample size of 82 would allow for power of 0.80. Although both intervention and contrast students' performances were still low at posttest, the performance of both groups improved approximately 10 standard score points over the course of the year. Analysis of the separate subtests of the oral language composite, however, yielded different patterns of performance; for listening comprehension and picture vocabulary, there were no differences between intervention and contrast students after adjusting for pretest performance (both $p > .05$ with small effect sizes—power was low, and per-group sample sizes required for power of .80 were high). However, there was a trend for intervention students to outperform contrast students on the verbal analogies subtest after adjusting for pretest performance, $F(1, 34) = 4.01$, $p < .053$, with a large effect size (+.77); in addition, power was .68, and a per-group sample size of only 27 would have allowed for power of .80.

Reading and academic achievement. On the WLPB-R word attack subtest, intervention students demonstrated a greater ability than contrast students to apply phonic and structural analysis skills to pronounce phonetically regular nonsense words in English after adjusting for pretest letter-word identification and PA composite performances, $F(1, 34) = 11.61$, $p < .002$, and the effect size of the difference between groups was large (+1.09). On the WLPB-R passage comprehension subtest, there was a significant difference between groups after adjusting for pretest letter-sound identification performance on this measure, $F(1, 36) = 13.82$, $p < .001$, with intervention students demonstrating greater ability to supply missing words to demonstrate comprehension in a cloze procedure; the effect size of the difference between groups was large (+1.08). A similar pattern was noted on the WLPB-R dictation subtest, $F(1, 37) = 5.81$, $p < .02$, with intervention students outperforming contrast students, with a strong effect size (+0.76). In addition, the performance of the intervention students on all three of these academic measures was within the average range at posttest.

Students were also administered two word reading fluency stories of the DIBELS (at levels gauged to correspond to beginning and end of first grade). As shown in Table 1, for students who provided data at pretest, mean performance on the DIBELS beginning-of-the-year story (words read correctly) was close to zero. However, at posttest, nearly all remaining students provided data, reading approximately 20 words correctly. For both English DIBELS stories, there were no significant differences on oral reading fluency between intervention and contrast students (both $p < .05$) after adjusting for pretest performance on rapid letter naming, with small, albeit posi-

TABLE 2. Postest Performance on Language and Reading Measures in English

Subtest/Group	*n*	*M*	*SD*	EFFECT	CI	*F*	*p*	GAIN*	*SD* DIFF
Letter naming:									
Letter name identification:[a]									
Intervention	22	24.50	2.6	+.59	−.05 to +1.23	1.72	*ns*	5.95	7.05
Contrast	18	21.83	6.0					6.94	5.74
Rapid letter naming:[b]									
Intervention	22	1.14	.3	+.88	+.23 to +1.53	9.68	.004	.72	.41
Contrast	18	.75	.5					.39	.34
Phonological processing:									
Letter-sound identification:[a]									
Intervention	22	24.36	2.2	+1.01	+.35 to +1.87	12.50	.001	7.91	5.64
Contrast	18	19.32	6.9					5.28	4.08
Nonword repetition:[a]									
Intervention	22	8.09	2.9	+.46	−.17 to 1.09	1.99	*ns*	3.45	3.85
Contrast	18	6.89	2.1					3.50	4.67
Phonemic awareness:									
Composite:[c]									
Intervention	22	60.12	10.4	+1.24	+.56 to +1.92	18.51	.000	29.16	9.5
Contrast	18	45.88	12.3					20.08	11.5
Language:									
Listening comprehension:[d]									
Intervention	21	68.52	18.3	+.26	−.37 to +.89	<1	*ns*	19.05	18.26
Contrast	18	63.78	17.1					16.61	16.73
Picture vocabulary:[d]									
Intervention	21	66.71	20.6	+.09	−.54 to +.72	1.32	*ns*	3.67	10.83
Contrast	18	64.83	19.4					9.83	13.78
Verbal analogies:[d]									
Intervention	21	90.14	11.2	+.77	+.10 to +1.45	4.01	.053	9.24	11.21
Contrast	16	82.25	8.0					7.50	11.53
Oral language composite:[d]									
Intervention	21	71.05	16.9	+.43	−.20 to +1.07	<1	*ns*	9.33	9.04
Contrast	18	63.33	18.1					10.00	12.78
Reading:									
Word attack:[d,e]									
Intervention	22	104.95	12.3	+1.09	+.40 to +1.78	11.61	.002	. . .	. . .
Contrast	16	91.19	12.4					. . .	. . .

Dictation (Woodcock):									
Intervention	22	93.77	16.7	+.76	+.11 to +1.42	5.81	.021	...	...
Contrast	17	80.71	16.9					...	...
Passage comprehension:[d,e]									
Intervention	22	98.23	13.8	+1.08	+.41 to +1.76	13.82	.001	...	...
Contrast	17	83.59	12.4					...	...
DIBELS BOY:[a,e]									
Intervention	22	21.68	14.5	+.16	−.51 to +.83	<1	*ns*	...	...
Contrast	14	18.57	23.8					...	...
DIBELS EOY:[a,e]									
Intervention	22	21.00	12.5	+.18	−.50 to +.85	<1	*ns*	...	...
Contrast	14	17.93	22.6					...	...

NOTE.—M = unadjusted mean of posttest performance; EFFECT = effect size (Cohen's d, adjusted for sample overestimation bias) of the difference between groups at posttest, based on unadjusted means; CI = confidence interval of adjusted effect size; F = statistical test of the posttest difference between adjusted (for pretest performance or other covariate) means of groups (degrees of freedom [df] are [1, $N - 3$] for all measures except dictation, where df are [1, $N - 2$], and word attack, where df are [1, $N - 4$] and N is INT + CON); $p = F$ test probability; GAIN = average change in performance from pretest (see Table 1) to posttest; DIBELS EOY = DIBELS End of Year Story. Dictation is a subtest of the WLPB-R and is unadjusted for pretest performance. Gain scores are not provided for subtests not given at both pre- and posttest (dictation and DIBELS EOY), or for measures where more data were available at posttest than pretest (word attack, passage comprehension, and DIBELS BOY) because the calculation of these gains would necessitate that the results be based on a sample different from all the other data provided for these measures in Table 2. All other abbreviations and notes are found in Table 1.

[a]Raw score.

[b]Letters per second.

[c]Raw score percent correct (from the sound matching, blending words, blending nonwords, segmenting words, and elision subtests of the CTOPP/TOPPS).

[d]Standard score ($M = 100$; $SD = 15$).

[e]Indicates measures for which missing data were present at pretest (due to inability), but available at posttest.

tive, effect sizes (+0.16 and +0.18). Power was low (0.07 and 0.08), with high per-group sample sizes required for power of 0.80, which may be due to the large variability in performance, particularly within the contrast group.

Posttest Performance: Spanish Measures

Results of posttest performance in Spanish are presented in Table 3. Included are means, effect sizes, significance tests, and gain scores for students who had test data at both times. For Spanish word attack, passage comprehension, and DIBELS (where many students did not provide data at pretest but did so at posttest), we followed a procedure similar to that described for the analogous English measures. For word attack, the PA composite was the covariate. For passage comprehension, the letter-sound identification and TOPP rapid letter naming subtests were the covariates. For DIBELS (beginning and end-of-year stories), the covariates were letter name and TOPP-S rapid letter naming subtests. In general, whereas the posttest performances of the intervention students across *English* outcome measures were consistently, significantly, and meaningfully greater than those of contrast students, we observed fewer differences between groups on *Spanish* outcome measures.

Letter naming. Within the domain of letter naming, there were no differences in performance at posttest after adjusting for pretest performance on either measure; in addition, effect sizes of the nonsignificant differences were generally small (with low power and also large per-group sample sizes needed to attain power of 0.80).

Phonological processing. Within the Spanish phonological processing domain, however, intervention students outperformed contrast students on the PA composite after adjusting for pretest performance, $F(1, 37) = 7.30$, $p < .01$; intervention students correctly answered an average of 58% of the items across the subtests administered relative to 46% for contrast students, and the standardized effect size of the difference between groups was large (+0.76). On other measures of phonological processing (letter-sound identification and nonword repetition), the groups did not differ significantly from one another, although effect sizes favoring intervention students were modest (+0.48 and +0.37), with low power (0.32 and 0.21) and per-group sample sizes of 68 and 113 (respectively) needed to attain power of 0.80.

Oral language. Within the Spanish oral language domain, there was a trend for the intervention students to outperform contrast students on the WLPB-R oral language composite, $F(1, 26) = 3.38$, $p < .07$, although the raw effect size (+0.01) was negligible. There was also a trend for contrast students to outperform intervention students on WLPB-R listening comprehension, $F(1, 36) = 3.93$, $p < .05$, although again the raw effect size (−0.04) was negligible. The reason for the discrepancies between the statistical significance and effect size results may be traced to the fact that effect sizes are based on posttest data only, whereas the statistical results take into consideration pretest performance as well. In the case of these measures, intervention students had higher pretest performance (though not statistically significant); over time, these students improved slightly, whereas contrast students improved considerably (e.g., more than a full standard deviation in the case of listening comprehension).

Reading and academic achievement. Intervention students outperformed contrast students on the WLPB-R word attack, $F(1, 37) = 7.86$, $p < .008$, and passage comprehension, $F(1, 34) = 4.23$, $p < .05$, subtests after adjusting for pretest performance on covariates. The effect sizes of the differences between groups were large for both measures (+0.87 and +0.81, respectively), though not as large as those for the analogous English measures. On the WLPB-R dictation subtest, there were no significant differences between intervention and contrast students, $F(1, 38) = 1.39$, $p > .05$. There

was a moderate effect size of +0.42 favoring intervention students, although power was low (0.26) and a sample size of 87 per group would be needed to attain power of 0.80. On the two DIBELS stories, there were also no significant differences between groups on the number of words read correctly (minus the number read incorrectly) after adjusting for pretest letter name identification and rapid letter naming performance (both $p > .05$). Effect sizes were small for both beginning (+0.31) and end-of-year (+0.25) stories, with low power and large sample sizes needed to attain power of 0.80.

Discussion

In this research we examined the effectiveness of a reading and oracy intervention provided in addition to core reading instruction for students struggling to read in English. We also examined how well reading skills in English influenced skills in students' native language (i.e., Spanish). Our results lead us to conclude that the ELL students who were provided this intervention program responded favorably for beginning reading skills, including comprehension. Gains made in Spanish (not the language of instruction for reading) were fewer and generally less strong than those observed in English, with the strongest differences favoring intervention students in the areas of phonological awareness and related reading skills as well as word attack and reading comprehension. It is important to note that these students were not proficient language speakers in English or Spanish (pretest scores in both languages averaged more than 2 standard deviations below the mean) and that students were disadvantaged in beginning reading skills in both languages.

Efficacy of Supplemental Intervention

The effects of the intervention may be better understood by recognizing that the contrast group received core instruction that was of high quality. Even so, students who participated in the supplemental intervention made statistically significantly greater gains across the year than students who participated in the core reading program with school-provided intervention on multiple measures of reading in English, including phonological access (RAN), phonemic awareness, letter knowledge (sounds), word attack, passage comprehension, and spelling dictation. These findings are educationally significant, given the close association of phonological awareness and letter knowledge (Wagner, 1988) to successful reading acquisition and development in later grades, and the importance of reading comprehension as the ultimate goal of reading instruction.

Effect sizes also indicate that, beyond achieving statistical significance, the differences between groups on most measures were moderate to large, and thus educationally relevant. Because the sample size was relatively small, our power to detect statistical significance was relatively low. Thus, examining effect size information is particularly important for understanding the true effects of the intervention. On measures for which statistical significance was achieved, effect sizes ranged from 0.76 to 1.24, with the largest effect being observed on the phonological awareness cluster. Most effect sizes on measures for which statistical significance was not detected favored the intervention group and were educationally meaningful (0.40 or above). In fact, only measures of listening comprehension, picture vocabulary, and oral reading fluency were below the 0.40 level. The average effect size for all nonsignificant measures was 0.37 (range = 0.09 to 0.77), and for measures with educationally meaningful effect sizes, the average effect size was 0.56. These effect sizes are more impressive when considering that they represent the magnitude of the value-added effect of the intervention beyond the effect of high-quality core reading instruction.

Generalization of Instructional Principles

This study indicates that EL readers with initial low literacy and language skills

TABLE 3. Posttest Performance on Language and Reading Measures in Spanish

Type/Subtest Name	*n*	*M*	*SD*	EFFECT	CI	*F*	*p*	GAIN*	*SD* DIFF
Letter naming measures:									
Letter name identification:[a]									
Intervention	22	16.18	7.4	+.13	−.50 to +.75	<1	*ns*	3.27	6.4
Contrast	18	15.17	8.5					4.39	6.0
Rapid letter naming:[b]									
Intervention	20	.46	.5	+.17	−.47 to +.80	<1	*ns*	.39	.4
Contrast	18	.38	.5					.27	.4
Phonological processing measures:									
Letter-sound identification:[a]									
Intervention	22	20.64	8.1	+.48	−.15 to +1.11	<1	*ns*	5.32	8.2
Contrast	18	16.94	7.0					5.22	4.4
Nonword repetition:[a]									
Intervention	22	9.64	2.8	+.37	−.26 to +1.00	<1	*ns*	1.14	2.7
Contrast	18	8.50	3.3					1.39	4.3
Phonemic awareness:									
Composite:[c]									
Intervention	22	58.32	15.2	+.76	+.11 to +1.40	7.30	.010	27.91	13.7
Contrast	18	45.75	17.4					17.98	10.4
Language measures:									
Listening comprehension:[d]									
Intervention	21	78.71	18.5	−.04	−.67 to +.59	3.93	.055	3.86	13.5
Contrast	18	79.56	24.6					17.33	18.6
Picture vocabulary:[d]									
Intervention	21	58.24	28.2	−.22	−.87 to .42	1.03	*ns*	.57	15.2
Contrast	17	65.12	32.3					5.47	14.6
Verbal analogies:[d]									
Intervention	21	89.33	11.0	+.19	−.44 to .82	<1	*ns*	4.71	11.9
Contrast	18	86.83	14.3					10.28	13.8
Oral language composite:[d]									
Intervention	21	71.81	19.3	+.01	−.62 to +.64	3.38	.074	3.90	9.6
Contrast	18	71.56	27.4					10.83	12.0

Reading measures:									
Word attack:[d,e]									
Intervention	22	90.68	20.7	+.87	+.21 to 1.52	7.86	.008	...	...
Contrast	18	73.11	18.9					...	...
Dictation:[d]									
Intervention	22	80.23	25.6	+.42	−.21 to +1.05	1.39	*ns*	...	...
Contrast	18	71.44	20.4					...	...
Passage comprehension:[d,e]									
Intervention	21	88.10	16.6	+.81	+.14 to +1.47	4.23	.05	...	...
Contrast	17	76.41	10.4					...	...
DIBELS BOY:[d,e]									
Intervention	21	7.90	6.8	+.31	−.35 to .98	1.50	*ns*	...	...
Contrast	15	5.80	6.8					...	...
DIBELS EOY:[d, e]									
Intervention	21	7.00	5.5	+.25	−.42 to +.91	<1	*ns*	...	...
Contrast	15	5.33	7.9					...	...

NOTE.—Word attack *df* are (1, $N - 3$), and *df* for passage comprehension, DIBELS BOY, and DIBELS EOY are (1, $N - 4$). See Tables 1 and 2 for all other notes and abbreviations.

[a]Raw score.

[b]Letters per second.

[c]Raw score percentage correct (from the sound matching, blending words, blending nonwords, segmenting words, and elision subtests of the CTOPP/TOPPS).

[d]Standard score ($M = 100$; $SD = 15$).

[e]Indicates measures for which missing data were present at present at pretest (due to inability), but available at posttest.

can profit from instruction that builds in complexity over time and that provides such features as reviews, practice, discussion with the teacher, and repeated reading and reading of text structures to improve comprehension. The findings are in line with outcomes observed with monolingual English speakers who were at risk for reading failure who received this same supplemental reading intervention without the oracy component (Mathes et al., 2004).

Reading Comprehension

Without question, the most important outcome is the influence of the intervention on reading comprehension. This study, as well as the previous study conducted with bilingual students in Spanish (Vaughn et al., 2006), reports findings that demonstrate both statistically and practically significant gains for reading comprehension. These findings can be contrasted with the study conducted with monolingual English-speaking students using the same English curriculum (Mathes et al., 2005) in which intervention students demonstrated no statistically significant gains in reading comprehension (ES = 0.21.). However, in the current study ELLs receiving supplemental instruction achieved significantly greater gains than ELLS in the contrast condition (ES = 1.08). We suspect that the addition of the story retell component may account for this difference, particularly because this component was not included in the study with monolingual English at-risk readers. The relatively high reading comprehension gains by participants in the study reported here can be understood by considering their practical significance. The intervention students made gains in English reading comprehension from more than 1 standard deviation below the normative sample to within the normal range (standard score = 98) at posttest. This contrasts with students in the control group, whose standard scores at pretest were also low (82) and who made relatively little gain in reading comprehension in English (standard score = 84).

Importance of Oracy Component and ESL Strategies

Both the intervention and contrast students demonstrated low oral language performance in both English and Spanish at both testing times—beginning and end of first grade. Both groups were performing on average more than 2 standard deviations below the normative sample on overall language ability in both languages. Though the intervention students made some significant gains over control students on language subtests, the rate of progress for both groups in both languages was very low and likely to influence reading and understanding of text as these students progress through the grades. This low language ability is no small consideration, and we are particularly interested in how these students progress in language and literacy as they proceed through the grades.

The higher language composite scores of intervention students are evident in particular in their ability to make language connection as reflected in their higher performance on verbal analogies. Likewise, their higher performance on passage comprehension may be the result of their ability to apply greater oral language skills to making meaning from text.

Influence of English Intervention on Spanish Literacy/Oracy

There is little information about how literacy instruction in one alphabetic language influences literacy acquisition in another alphabetic language, particularly for at-risk readers who are acquiring literacy skills in their second language and have limited literacy skills in their first language. In examining relative progress in both English and Spanish, we found that gains in English appear to influence gains in Spanish. In particular, students who participated in the intervention, as compared with contrast students, demonstrated significantly greater gains in Spanish phonemic awareness, word attack, and passage comprehen-

sion and showed a trend toward stronger listening comprehension. It is important to note that the ELLs who participated in this study demonstrated poor language skills in both Spanish and English. Thus, additional transfer may occur from English to Spanish in students who have stronger beginning language skills. As children become more proficient readers, they may be better able to transfer what they have learned about reading in English to reading in Spanish. However, even in the present sample, transfer was observed in some of the most relevant domains, including phonological awareness and passage comprehension, suggesting that children were attentive to the phonemes of words and could generate meaning from text in both their native language and in English. We will be following this cohort of children as they matriculate across grades and will be able to report more fully on transfer between languages with this population in the future.

Future Research

Clearly there is need for continued research with ELLs who are struggling readers. A primary question that remains to be answered concerns the long-term outcome for the ELLs who participated in this research. To address this question, our team is following this cohort of children into their third-grade year. Relatedly, we need to determine if the results of the current research are replicable. Currently, we are implementing the supplemental intervention with a second cohort of first-grade ELLs who are struggling readers. Thus we will soon have data about the academic trajectories of students who did and did not receive supplemental interventions in first grade and whether or not the outcomes in this study are replicable.

Still, many questions remain unanswered. Intervention students had low initial language proficiency in both their native and second languages but moved within the average range in reading in their second language. However, these children may need supplemental instruction in the future. There is little guidance about how much supplemental instruction will be necessary to assure reading success throughout the grades.

Another issue that remains unanswered from the current research is the value of each intervention component. Because the reading program, the ESL support activities, and the retell routine were conducted as a package, disentangling effects of components was not possible. In particular, we are interested in knowing effects of the retell routine on oracy and comprehension because this component can easily be conducted independently.

Perhaps the most pertinent question not addressed in this study concerns the language of instruction. We provided instruction in English because these children were receiving core instruction in English. We have also conducted a second study with a similar design in which ELLs with similar demographics and language proficiency received instruction in Spanish (Vaughn et al., 2006). Because we are following both groups of students until third grade, we will better understand the influence of early intervention and language instruction on reading performance in English and Spanish.

Conclusion

Although there is a need for continued research with ELLs who are struggling readers, this study answers several important and often-debated questions. First and foremost, it affirms that ELLs at risk for reading problems and provided intervention in English increased scores considerably and on par with their monolingual English peers on most subtests and performed better on reading comprehension (Mathes et al., 2004). Likewise, ELLs learning to read in English make significant growth when this supplemental instruction includes phonemic awareness, letter knowledge, alphabetic decoding, decodable text practice, and comprehension strategies. We speculate

that it may be necessary to include more vocabulary and prior knowledge work to improve word knowledge and continue to influence reading comprehension. Of course, the supplemental instruction provided in this research also supported the learning of Ells by incorporating ESL strategies. Thus, we have demonstrated that ELLs make substantially greater progress in learning to read in English when instruction incorporates the same critical content shown to be effective with English speakers but also provides language support using ESL strategies. Our outcomes also suggest that ELLs have stronger outcomes when they engage in the type of language listening and discussion provided in the retell routine. What remains to be seen is if monolingual English readers will make similar gains if they too participate in retell. Similarly, future research will show if these students can transfer what they have learned about reading in English to their native language. Even so, it is clear that ELLs who are struggling readers profit from supplemental instruction. The goal now is how to ensure that those who need supplemental instruction receive it.

Note

This research was funded by a grant from the National Institute of Child Health and Human Development and the Institute of Educational Sciences (grant award PO1 HD 39521, Development of English Literacy in Spanish-Speaking Children). We would like to recognize the excellent contribution of Russell Gersten, whose editorial suggestions and advice improved considerably the quality of this article. We would also like to thank Jack Fletcher and Barbara Foorman, whose outstanding support for the program project positively influenced the design and data collection of this study.

References

Araujo, L. (2002). The literacy development of kindergarten English-language learners. *Journal of Research In Childhood Education*, **16**, 232–247.

Beck, I. L., McKeown, M. G., & Kucan, L. (2002). *Bringing words to life: Robust vocabulary instruction; solving problems in the teaching of literacy*. New York: Guilford.

Borenstein, M., Rothstein, H., Cohen, J., Schoenfeld, D., & Berlin, J. (2001). *Power and precision, verson 2.1: A computer program for statistical power analysis and confidence*. Mahwah, NJ: Erlbaum.

Branum-Martin, L., Metha, P., Fletcher, J. M., Carlson, C. D., Ortiz, A., Mario, C., & Francis, D. J. (2006). Bilingual phonological awareness: Multilevel construct validation among Spanish-speaking kindergartners in transitional bilingual education classrooms. *Journal of Educational Psychology*, **98**(1), 170–181.

Carnine, D W., Silbert, J., & Kame'enui, E. J. (1997). *Direct instruction reading* (3d ed.). Upper Saddle River, NJ: Merrill/Prentice-Hall.

Carreker, S. (1999). *Language enrichment*. Bellaire, TX: Neuhaus Education Center.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2d ed.). Hillsdale, NJ: Erlbaum.

Denton, C. A., Anthony, J. L., Parker, R., & Hasbrouck, J. (2004). Effects of two tutoring programs on the English reading development of Spanish-English bilingual students. *Elementary School Journal*, **104**(4), 289–305.

Edelsky, C. (1986). *Writing in a bilingual program: Había una vez*. Norwood, NJ: Ablex.

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**(3), 393–447.

Ehri, L. C., & Wilce, L. S. (1983). Development of word identification speed in skilled and less skilled beginning readers. *Journal of Educational Psychology*, **75**, 3–18.

Elley, W. B., & Magubhai, F. (1983). The impact of reading on second-language learning. *Reading Research Quartely*, **19**, 53–67.

Fitzgerald, J. (1995a). English-as-a-second language learners' cognitive reading processes: A review of research in the United States. *Review of Educational Research*, **65**, 145–190.

Fitzgerald, J. (1995b). English-as-a-second-language reading instruction in the United States: A research review. *Journal of Reading Behavior*, **27**, 115–152.

Fitzgerald, J., & Noblit, G. W. (1999). About hopes, aspirations, and uncertainty: First-grade English-language learners' emergent reading. *Journal of Literacy Research*, **31**, 133–182.

Fletcher, J. M., & Lyon, G. R. (1998). Reading: A

research-based approach. In W. M. Evers (Ed.), *What's gone wrong in America's classrooms?* (pp. 49–90). Stanford, CA: Hoover Institute.

Foorman, B. R., Goldenberg, C., Carlson, C., Saunders, W., & Pollard-Durodola, S. (2004). How teachers allocate time during literacy instruction in primary grade bilingual classrooms. In M. McCardle & V. Chhabra (Eds.), *The voice of evidence in reading research* (pp. 289–328). Baltimore: Paul H. Brookes

Foorman, B. R., & Schatschneider, C. (2003). Measurement of teaching practices during reading/language arts instruction and its relationship to student achievement. In S. Vaughn & K. L. Briggs (Eds.), *Reading in the classroom: Systems for the observation of teaching and learning* (pp. 1–30). Baltimore: Paul H. Brookes.

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.

Gersten, R., & Baker, S. (2000). What we know about effective instructional practices for English-language learners. *Exceptional Children,* **66**, 454–470.

Gersten, R., & Geva, E. (2003). Teaching reading to early language learners. *Educational Leadership,* **60**, 44–49.

Geva, E., & Wang, M. (2001). The development of basic reading skills in children: A cross-language perspective. *Annual Review of Applied Linguistics,* **21**, 181–204.

Geva, E., Yaghoub-Zadeh, Z., & Schuster, B. (2000). Understanding individual differences in word recognition skills of ESL children. *Annals of Dyslexia,* **50**, 123–154.

Good, R. H., Bank, N., & Watson, J. M. (Eds.). (2003). *Indicadores dinámicos del exito en la lectura.* Eugene, OR: Institute for the Development of Educational Achievement.

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

Goswami, U. (1993). Toward an interactive analogy model of reading development: Decoding vowel graphemes in beginning reading. *Journal of Experimental Child Psychology,* **56**, 443–475.

Grek, M. L., Mathes, P. G., & Torgesen, J. K. (2003). Similarities and differences between experienced teachers and trained paraprofessionals: An observational analysis. In S. Vaughn & K. L. Briggs (Eds.), *Reading in the classroom: Systems for the observation of teaching and learning* (pp. 267–296). Baltimore: Paul H. Brookes.

Gunn, B. K., Biglan, A., Smolkowski, K., & Ary, D. V. (2000). The efficacy of supplemental instruction in decoding skills for Hispanic and non-Hispanic students in early elementary school. *Journal of Special Education,* **34**(2), 90.

Hasbrouck, J. E., & Tindal, G. (1992). Curriculum-based oral reading fluency norms for students in grades 2 through 5. *Teaching Exceptional Children*, **24**(3), 41–44.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis.* New York: Academic Press.

Kuhn, M. R., & Stahl, S. A. (2003). Fluency: A review of developmental and remedial practices. *Journal of Educational Psychology*, **95**(1), 3–21.

Linan-Thompson, S., Vaughn, S., Hickman-Davis, P., & Kouzekanani, K. (2003). Effectiveness of supplemental reading instruction of English language learners with reading difficulties. *Elementary School Journal*, **103**(3), 221–238.

Lopez, G. (2004). Bringing the mountain to Mohammed: Parent involvement in migrant-impacted schools. In *Scholars in the field: The challenges of migrant education* (pp. 134–147). Charleston, WV: Migrant Education Program.

Mathes, P. G., Denton, C. A., Fletcher, J. M., Anthony, J. L., Francis, D. J., & Schatschneider, C. (2005). An evaluation of two reading interventions derived from diverse models. *Reading Research Quarterly,* **40**, 148–183.

Mathes, P. G., Torgesen, J. K., Wahl, M., Menchetti, J. C., & Grek, M. L. (2004). *Proactive beginning reading: Intensive small-group instruction for struggling readers.* Curriculum developed with funds provided by the National Institute of Child Health and Human Development (No. R01 HD), Prevention and Remediation of Reading Disabilities, Washington, DC.

McGraw-Hill Reading. (2001). New York: McGraw-Hill School Division.

National Reading Panel. (2000). *Report of the national reading panel: Teaching children to read; an evidence-based assessment of the scientific research literature on reading and its implications for reading instruction* (NIH Publication No. 00–4769). Washington, DC: U.S. Government Printing Office.

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

Paulesu, E., Demonet, J.-F., Fazio, F., McCrory, E., Chanoine, V., Brunswick, N., et al. (2001).

Dyslexia: Cultural diversity and biological unity. *Science*, **291**, 2165–2167.

Pressley, M. (1998). *Reading instruction that works: The case for balanced teaching*. New York: Guilford.

Rayner, K., Foorman, B. R., Perfetti, C. A., Pesetsky, D., & Seidenberg, M. S. (2001). How psychological science informs the teaching of reading [Monograph]. *Psychology Science: A Journal of the American Psychological Society*, **2**, 31–74.

Schatschneider, C., Francis, D., Foorman, B., Fletcher, J., & Mehta, D. (1999). The dimensionality of phonological awareness: An applicaton of item response theory. *Journal of Educational Psychology*, **91**, 439–449.

Sebastian, N., Cuetos, F., Marti, M. A., & Carreiras, M. F. (2000). *LEXESP: Léxico informatizado del español* (edición en CD-ROM). Barcelona: Edicions de la Universitat de Barcelona.

Signorini, A. (1997). Word reading in Spanish: A comparison between skilled and less skilled beginning readers. *Applied Psycholinguistics*, **18**, 319–344.

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

Torgesen, J. K., Mathes, P. M., & Grek, M. L. (2002, February). *Effectiveness of an early intervention curriculum that is closely coordinated with the regular classroom reading curriculum.* Paper presented at the Pacific Coast Research Conference, San Diego.

Treiman, R. (1984). Individual differences among children in spelling and reading styles. *Journal of Experimental Child Psychology*, **37**, 463–477.

Ulanoff, S. H., & Pucci, S. L. (1999). Learning words from books: The effects of read-aloud on second-language acquisition. *Bilingual Research Journal*, **23**, 409–422.

Vaughn, S., Cirino, P. T., Linan-Thompson, S., Mathes, P. G., Carlson, C. D., Cardenas-Hagan, E., et al. (in press). Effectiveness of a Spanish intervention and an English intervention for English language learners at risk for reading problems. *American Educational Research Journal.*

Vaughn, S., Linan-Thompson, S., Mathes, P. G., Cirino, P. T., Carlson, C. D., Pollard-Durodola, S. D., et al. (2006). Effectiveness of a Spanish intervention for first-grade English language learners at risk for reading difficulties. *Journal of Learning Disabilities*, **39**(1), 56–73.

Vaughn, S., Mathes, P., Linan-Thompson, S., & Francis, D. (2005). Teaching English language learners at risk for reading difficulties to read in Spanish or English. *Learning Disabilities Research and Practice*, **20**(1), 58–67.

Wagner, R. K. (1988). Causal relations between the development of phonological processing abilities and the acquistion of reading skills: A meta-analysis. *Merrill-Palmer Quarterly*, **34**, 261–279.

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

Woodcock, R. W. (1991). *Woodcock Language Proficiency Battery—Revised*. Chicago: Riverside.

Woodcock, R. W., & Johnson, M. B. (1989). *Woodcock-Johnson Psycho-Educational Battery—Revised.* Chicago: Riverside.

Woodcock, R. W., & Munoz-Sandoval, A. F. (1993). *Language survey: Comprehensive manual*. Chicago: Riverside.

Zeno, S. M., Iyens, S. H., Millard, R. T., & Duvvuri, R. (1995). *The educator's word frequency guide.* New York: Touchstone Applied Science Associates.