

Supplemental Instruction in Decoding Skills for Hispanic and Non-Hispanic Students in Early Elementary School:

A Follow-Up

Barbara Gunn, Keith Smolkowski, Anthony Biglan, and Carol Black,
Oregon Research Institute

This article describes a follow-up study that experimentally evaluated the effects of supplemental reading instruction for children in kindergarten through Grade 3. Students from 10 elementary schools in three school districts were screened using the *Dynamic Indicators of Basic Early Literacy Skills*. Two hundred fifty-six K-2 students were identified for participation, then randomly assigned to receive or not receive 2 years of supplemental reading instruction that taught basic decoding and comprehension skills. Reading ability was measured in the fall prior to the first year of the intervention and again in the spring of Years 1, 2, and 3. At the end of the 2-year intervention, children who received the supplemental instruction performed better on measures of word attack, word identification, oral reading fluency, vocabulary, and reading comprehension. One year after the intervention, children in the supplemental instruction group still showed greater improvement in word attack and oral reading fluency than the comparison students.

The vital importance of acquiring literacy skills in the early elementary grades is well established (e.g., Cunningham & Stanovich, 1998; Juel, 1988; Slavin et al., 1996). Children who fall behind their classmates in learning to read not only have to catch up but also must use their limited reading ability to keep pace with the daily introduction of new lessons and skills. This is no easy task. Students who struggle with reading are especially vulnerable to academic underachievement, and studies show that a large number of poor readers at the end of first grade remain poor readers in third grade and beyond (Francis, Shaywitz, Stuebing, Shaywitz, & Fletcher, 1996; Juel, 1988; Slavin et al., 1996). Those who fail to catch up are frequently referred for remedial or special education services and join the growing population of students identified with deficits in reading and language (Greenwood Institute, 1994).

Reading difficulty among Spanish-speaking children is a particular concern. On average, Spanish-speaking students in the United States have lower levels of reading achievement in English than other groups (Fitzgerald, 1995) and are at high risk for academic failure as a result. Moreover, many Spanish-speaking students trail behind their classmates academically

throughout elementary school and are referred in disproportionately high numbers for special education (President's Advisory Commission, 1996). Between the years of 1976 and 1994, Hispanic children with learning disabilities increased from 24% to 51% among all students with learning disabilities.

The Need for Supplemental Instruction in Reading

Research on beginning reading has identified the processes underlying the acquisition of reading skill (e.g., Ehri, 1998; Wagner et al., 1997) and highlighted the essential features of beginning reading instruction (e.g., Adams, 1990; Kameenui et al., 1998; Snow, Burns, & Griffin, 1998). However, practical difficulties often stand in the way of providing optimal instruction. Increasing numbers of states have curriculum standards that align with research-based principles of reading instruction. Yet, schools and districts still use widely disparate methods for teaching reading (Stotsky, 2000). Teachers, experienced as well as novice, do not receive the training or ma-

terials they need to effectively teach reading skills (Moats, 1995), and their task is complicated by non-English-speaking children, families' high mobility rates, and the lack of literacy preparation among many early elementary school children.

It is therefore appropriate to explore methods for ensuring that children at risk for reading disability develop the reading skills that allow them to succeed academically. This study describes the effects of supplemental instruction provided by trained instructional assistants. Supplemental instructional approaches have been advocated as a means by which non-school organizations can increase reading achievement and foster literacy development for all students (U.S. Department of Education, 1998), and such instruction may be of particular value for Hispanic students. However, empirical evidence regarding the value of such instruction is currently lacking.

We (Gunn, Biglan, Smolkowski, & Ary, 2000) reported the initial effects of 2 years of supplemental instruction for children in Grades K through 3 who were at risk for reading disability. Both Hispanic and non-Hispanic children were included in the study. Ethnicity was self-reported by parents. *Hispanic* was defined as Mexican American; *non-Hispanic* as European American. Children who received supplemental instruction improved significantly more than their controls on a measure of word attack skills after 1 year of instruction. The treatment and control students did not differ significantly on measures of letter-word identification or oral reading fluency, although all differences favored the intervention condition. At the end of the second year of intervention, children who received supplemental instruction improved significantly more on the measures of letter-word identification, word attack, and oral reading fluency than the children who not receive the instruction. Instructed children also performed better on reading vocabulary and passage comprehension measures that were included with the third assessment. These findings held for Hispanic and non-Hispanic children alike. Indeed, instruction produced significant effects on oral reading fluency for children who spoke only Spanish at the outset of instruction. Findings from the earlier study (Gunn et al., 2000) supported the immediate benefits of supplemental reading instruction for Hispanic and non-Hispanic students in K through Grade 3. The purpose of the current study was to examine whether the benefits of supplemental instruction were maintained 1 year after the intervention concluded.

Method

Data for the study come from an experimental evaluation of a program to prevent academic and social failure among K-3 students. The Schools and Homes in Partnership (SHIP) project was conducted in school districts in three small Oregon communities, each of which has a large Mexican American population. This article reports on the effects of the reading intervention on reading outcomes 1 year following the end of the intervention. We do not describe effects on social behav-

ior because only a subsample of children screened for aggressive social behavior received reading instruction.

Procedure

We screened Hispanic and non-Hispanic students in K through third grade on measures of aggressive social behavior and reading (or prereading) skill. We recruited the children and their families to participate in the study and conducted pretest assessments of the children's reading skill and behavior. Students were then matched on ethnic group and randomized according to their scores on the behavioral measure or the reading pretest. All students in the treatment condition who had scores below their grade-level equivalent on at least two reading pretests were eligible to receive the supplemental reading instruction. Participating students received 4 to 5 months of instruction in Year 1 and 9 months of instruction in Year 2. Students were tested on measures of reading skill at the end of each year of intervention, and again 1 year after the intervention concluded.

Screening. Students in K, first grade, and second grade were screened in the spring of the year before intervention. In the fall of the intervention year, students in the new K class were screened, along with students who had transferred into Grades 1, 2, or 3 in participating schools. All screening was done in English; however, bilingual assessors provided testing information in English and Spanish as needed, to make sure that students understood the test directions.

Assessments. Pre-intervention assessments (Time 1) were conducted in the fall for all children who met screening criteria, and a second assessment (Time 2) was given in the spring of that same academic year. Participating children were tested again in the spring of the following 2 years (Times 3 and 4). The Time 3 assessment was conducted at the end of the 2-year intervention, and the Time 4 assessment constituted a 1-year follow-up.

Randomization. All students who scored above 3.0 on the *Walker-McConnell Test of Social Skills* (Walker & McConnell, 1988) were grouped by grade and ethnicity (Hispanic or non-Hispanic); rank ordered by their reading ability in English; matched on ethnicity; and, beginning with the poorest pair of readers, randomly assigned to conditions. In this process, then, we randomly placed one student in each pair into the treatment condition and the other into the control condition. The remaining students were matched into pairs by their scores on the *Walker-McConnell* and similarly randomly assigned to conditions.

Participants

School districts in three small Oregon communities participated in this study. According to 1990 census data, their pop-

ulations were 4,632 in Community A, 13,559 in B, and 13,404 in C, and the proportion of Hispanics in the communities was 10.5%, 9.2%, and 31.5%, respectively. The school districts in these communities served a wider population than the cities in which they were based. In the fall of 1997, the proportion of Hispanics enrolled in each school district was 30.9% in A, 25.0% in B, and 59.7% in C. There were 1, 5, and 4 participating schools in Communities A, B, and C, respectively.

Two hundred fifty-six students who were below grade level in reading participated in the evaluation of supplemental instruction. This was a subsample of the 284 students who qualified for participation in the study on the basis of either their reading skill scores or their social behavior scores. This subsample comprised 156 students who were screened into the study on the basis of poor reading skills, and 100 students who were selected on the basis of aggressive social behavior but were also below grade level in reading skill. To qualify for supplemental instruction, students selected by the aggression criterion also had to score below grade level on two or more of the baseline measures of reading skill. Baseline measures included oral reading fluency on grade-level passages and the *Woodcock-Johnson Revised Tests of Achievement* Letter-Word Identification and Word Attack subtests. Grade equivalents were used as the criteria for eligibility for supplemental reading instruction for the *Woodcock* because it was normed on a sample of students across a wide grade range and includes test items that are distributed uniformly over a range of difficulty (Woodcock & Mather, 1989, 1990). The test reflects the level of task difficulty a student can attain, and grade equivalents are recommended as appropriate for instructional planning. However, grade equivalents were not used in any of the subsequent analyses of student performance.

One hundred forty-two of the participants were boys and 114 were girls. Sixty-two percent of the students were Hispanic and 38% were non-Hispanic. Parents provided information about their ethnic identity, nativity, and language use. These data indicated that 94% of the Hispanic sample were of Mexican heritage, 5% were from Central America, and the remainder were from some other Latin American country. About 9% were born in the United States; 85%, in Mexico; and the remainder, in another Latin American country. Eighty-four percent spoke only or mostly Spanish. Table 1 presents the breakdown of participants by ethnicity, gender, and grade level, and Table 2 presents the breakdown of participants by condition.

Screening Measures

Teachers' Ratings of Aggressive Behavior. The aggression scale of the Teacher Rating Form (TRF) of the *Child Behavior Check List* (Achenbach, 1991) was used to screen for aggressive behavior. Each child's classroom teacher rated 25 items on 3-point scales to indicate how characteristic a behavior (e.g., disrupts class discipline) was of the student during the past 2 months. Wilson and Bullock (1989) cited a .89

TABLE 1. Characteristics of Participants by Ethnicity, Gender, Condition, and Grade Level

Grade level		Basis for selection		
		Aggression	Reading	Total
Boys				
Kindergarten	Hispanic	0	13	13
	Non-Hispanic	8	7	15
	Total	8	20	28
1st grade	Hispanic	8	19	27
	Non-Hispanic	13	3	16
	Total	20	22	42
2nd grade	Hispanic	10	16	26
	Non-Hispanic	14	4	18
	Total	24	20	44
3rd grade	Hispanic	5	14	19
	Non-Hispanic	3	6	9
	Total	8	20	28
Girls				
Kindergarten	Hispanic	3	12	15
	Non-Hispanic	4	4	8
	Total	7	16	23
1st grade	Hispanic	7	21	28
	Non-Hispanic	5	5	10
	Total	12	26	38
2nd grade	Hispanic	5	8	13
	Non-Hispanic	13	4	17
	Total	18	12	30
3rd grade	Hispanic	0	17	17
	Non-Hispanic	3	3	6
	Total	3	20	23

test-retest reliability for the teacher ratings and reported that the measure correctly identified 76% of children referred for emotional/behavior problems from nonreferred children.

Dynamic Indicators of Basic Early Literacy Skills (DIBELS). The DIBELS (Good, Kaminski, Laimon, & Johnson, 1992) Letter Naming Fluency (LNF) task measures the speed and accuracy with which students identify letter names during a 1-minute timing. Rapid-naming tasks approximate the requirements in reading running text (Wolf & Bowers, 1999). The LNF score is reported as the number of letters named correctly per minute.

Phonemic awareness is a strong predictor of early reading success (Wagner, Torgesen, Laughon, Simmons, & Rashotte, 1993). The DIBELS Phoneme Segmentation Fluency task assesses children's ability to break apart the sounds in words. It is a timed 10-word measure in which the examiner says a word, then asks the student to segment the word into phonemes. The time taken to segment the words and the total number of correct segments per minute are reported as the rate of correct segments per minute.

TABLE 2. Characteristics of Participants by Condition

Variable		Ethnicity		All participants
		Non-Hispanic	Hispanic	
Age				
Control	<i>M</i>	7.42	7.09	7.21
	<i>SD</i>	.97	.97	.98
	<i>n</i>	35	61	96
Intervention	<i>M</i>	6.94	7.37	7.19
	<i>SD</i>	1.02	1.03	1.04
	<i>n</i>	39	59	98
Grade				
Control	<i>M</i>	1.67	1.41	1.50
	<i>SD</i>	.97	1.01	.99
	<i>n</i>	35	61	96
Intervention	<i>M</i>	1.18	1.75	1.53
	<i>SD</i>	1.05	1.05	1.08
	<i>n</i>	39	60	99
Male				
Control		63%	59%	60%
Intervention		51%	53%	53%
Selected for aggressive behavior				
Control		66%	34%	46%
Intervention		62%	22%	37%

Note. Age and grade are reported as the average at Time 1.

The DIBELS Phoneme Onset Fluency task is a simpler measure of phonological awareness than phoneme segmentation and was used to screen only the K children. On 12 of the 16 items, children are shown sets of four pictures of familiar objects and are asked to point to the picture that begins with the target sound. On the remaining four items, children are asked to produce the initial sound of the target word (e.g., "What sound does 'pumpkin' begin with?"). The time taken to identify the picture or sound and the total number of correct responses are reported as the rate of correct segments per minute. Good and Kaminski (1996) provided reliability and validity data for all of the DIBELS measures.

Oral reading fluency typically is identified as the number of words read correctly from text in 1 minute (Deno, 1985). This measure was used for screening students in Grades 1 and 2 (and also in the T1-T3 assessments for students in K through fourth grade, described later). First-grade students read two grade-level passages, and second-grade students read three grade-level passages. They were timed to determine how many words they could correctly read in 1 minute. The mean score for the 1-minute timings was recorded and included in the screening composite score used to identify students for

participation. Marston (1989) reported reliability and validity data for the reading fluency assessment. Hintze, Owen, Shapiro, and Daly (2000) reported that the dependability of three curriculum-based measurement reading passages for identifying reading problems and estimating performance discrepancies between individuals is about .95. On the basis of these data we hypothesize that the mean of two passages was dependable for individual screening purposes.

Screening Composite Score. A composite score was created by averaging the raw scores on the screening measures. In K the screening measures were rapid letter naming, phoneme onset fluency, and phoneme segmentation fluency. In Grade 1 the measures were phoneme segmentation fluency and two 1-minute samples of oral reading fluency, and in Grades 2 and 3 the measures were three 1-minute samples of oral reading fluency. The composite score was the average of individual student scores on all measures. In the first community where the study was conducted, children in the bottom 10% of the distribution on reading screening scores were selected; in the two subsequent communities, the children in the bottom 5% of the distribution were selected.

Outcome Measures of Reading Achievement

Participants' reading achievement was measured in the fall of the first year of the study and at all three subsequent assessments. Testing was conducted in English for all students, as it was English language reading proficiency that instruction was intended to achieve. It might be noted in this regard that Baker and Good (1994) found that criterion measures of English reading and language proficiency, as predictors, were as reliable and valid for bilingual students as they were for English-only students and were sensitive to the reading progress of bilingual students.

Oral Reading Fluency. A means of measuring reading proficiency is to obtain repeated samples of students' reading fluency on grade-level passages (Markell & Deno, 1997). We obtained three 1-minute reading samples from all students in Grades 1 through 3 as part of the annual assessment, using the same procedures used for the initial screening.

Woodcock-Johnson Revised Tests of Achievement (WJ-R ACH). Four subtests of the *Woodcock-Johnson Revised Tests of Achievement* were used to assess reading skill. The Letter-Word Identification and Word Attack subtests of the WJ-R ACH provide a Basic Reading cluster score that measures the ability to read irregular words and to use phonic and structural analysis. The Passage Comprehension and Reading Vocabulary subtests provide a Reading Comprehension cluster score. On the vocabulary subtest, students read words and supply appropriate meanings. For the comprehension subtest, students read passages and identify key words

that are missing. The Letter-Word Identification and Word Attack subtests were administered at each time point; the comprehension and vocabulary subtests were administered at the third and fourth assessments. Reliability and validity data for the Word Recognition and Attack subtests were reported by Woodcock and Mather (1989-1990).

Outcome Measures of Behavior

Teachers rated students on the *Walker-McConnell Scale of Social Competence and School Adjustment* (WM; Walker & McConnell, 1988). The scale consists of 19 items that refer to school behaviors (e.g., "does seat work assignments as directed") and interpersonal behaviors (e.g., "invites peers to play or share activities"). Psychometric data on the *Walker-McConnell* were collected using teachers and students from the four major census zones in the United States across 15 states, and the data are representative of students from a cross section of cultural, linguistic, and economic backgrounds. Reliability, validity, and normative data are available on the WM (Walker & McConnell, 1993).

Supplemental Reading Instruction

All participating students received daily reading instruction from their classroom teacher. Because students were randomly assigned within school, treatment and control students were in the same classrooms and received the same instruction. Treatment students also received 30 minutes of supplemental instruction daily for 4 to 5 months in Year 1 and 9 months in Year 2 from trained instructional assistants. This includes all children who were screened into the study on the basis of low reading skill and those children who were screened into the study on the basis of aggressive social behavior and were also below grade level in reading skill.

Because of the time needed to conduct assessments in the fall of the first year, the amount of time available for supplemental instruction ranged from 4 to 5 months. All students received a full year (i.e., 9 months) of reading instruction in the second year of the intervention. During the summer between Years 1 and 2, 79 students (39 girls and 45 boys) also attended summer school and received 30 minutes of reading instruction, 3 days a week for 5 weeks.

Research in the past 20 years has contributed to a model of the cognitive processes involved in reading acquisition (Adams, 1990; Anderson, Hiebert, Scott, & Wilkinson, 1985; Snow et al., 1998; Stanovich, 1986) and has identified the components of effective initial reading instruction. Reading Mastery (Engelmann & Bruner, 1988) and Corrective Reading (Engelmann, Carnine, & Johnson, 1988) were used for the supplemental instruction because they incorporate a large variety of empirically supported instructional techniques and have been extensively validated in small-group and whole class settings (Adams & Engelmann, 1996; Stahl & Miller, 1989).

Students were tested and placed in Reading Mastery if they were beginning readers in first or second grade. Reading Mastery provides explicit instruction in phonemic awareness, sound-letter correspondence, and blending. Sound-letter correspondences are selected and sequenced carefully to avoid confusion and to help students completely learn the most frequently used sounds. Students read code-based passages to practice new sound-letter correspondences and gradually build accuracy and fluency. Third- and fourth-grade students who had received beginning reading instruction but were still nonreaders or were reading below grade level were tested and placed in Corrective Reading. These students also received explicit instruction in phonemic awareness, sound-letter correspondence, and blending; however, new sounds were introduced at a slightly faster rate and the stories were geared to the interests of older children. In both programs, most students received instruction in groups of two to three. In cases where this was not feasible, one-to-one instruction was provided.

Nine instructional assistants were hired from project communities. We found that project employees who worked in their own communities were a strong asset. They understood their community norms, related well to participants, and were trusted as community members. Three assistants were certified teachers with 5 to 7 years of teaching experience. Seven assistants had experience working with elementary-school children in small-group or tutorial settings, and two assistants spoke Spanish and English. The assistants received 10 hours of pre-service training in procedures for testing and grouping students for instruction, presenting lessons, signaling for student responses, correcting errors, and motivating students with clear expectations and feedback. The theoretical framework underpinning effective reading instruction was also described in the training.

The instructional assistants were observed weekly during the first month of supplemental instruction and twice a month thereafter, to document fidelity of implementation. Using a direct observation checklist, a trained supervisor tallied the rate of student responses and the accuracy of their signaling and correction procedures. The instructional assistants were given direct feedback on their instruction, and, when needed, they observed a supervisor teach a lesson. Supervisors and assistants also consulted on ways to identify and clarify unfamiliar words and terms for the children as the need arose during lessons. Assistants met as a group twice a month with the trainer to practice and refine teaching techniques and discuss problems with individual students.

Results

Attrition

Results are reported for 195 participants. Of the 256 who were eligible for supplemental instruction and were randomized to condition, 9 (3.5%) had no Time 1 data and 52 (20.3%) had

no Time 4 data. Of the 52 with no reading data at Time 4, 36 had dropped out of the study because they moved out of participating school districts. The remaining 16 cases were not assessed either because the children frequently missed school or because the families were out of town for an extended period of time.

The effects of attrition were analyzed with 2×2 analyses of variance (ANOVAs) on the three reading measures that were available at Time 1 (Biglan et al., 1991). The independent variables in these analyses were instructional condition and attrition status (data missing at Time 4 vs. data not missing). The three measures were letter-word identification, word attack, and oral reading fluency. We had an interaction with condition for letter-word identification, $F(1, 230) = 8.09$, $p < .01$, $d = .36$, but not for word attack, $F(1, 230) = 1.39$, ns , $d = .16$, or oral reading fluency, $F(1, 231) = .858$, ns , $d = .12$. As discussed below, letter-word identification was not significant in the overall analyses. This may be partially a function of differential attrition or an indication that this measure was no longer the most sensitive measure of participants' reading ability at Time 4. The lack of interactions between attrition and intervention condition for word attack and oral reading fluency indicates that the internal validity of the study was not affected by attrition for these measures, as the two intervention conditions did not differ in the characteristics of those who were missing at Time 4.

Overview of Analyses

We began by conducting 2×2 analyses of variance with intervention versus control and Hispanic versus non-Hispanic as the independent variables. We chose ANOVAs of change scores (i.e., gain scores or difference scores) as our analysis method for all posttest measures for which we had an equivalent pretest. The dependent variables, then, for letter-word identification, word attack, and oral reading fluency were change scores between Time 1 and Time 4. Because the Vocabulary and Comprehension subtests of the *Woodcock* were not administered at Time 1, it was necessary to conduct an analysis of covariance (ANCOVA) on these variables, using oral reading fluency at Time 1 as the covariate. Skewness and kurtosis were generally low, and examination of normal and detrended normal probability plots indicated acceptable distributions of scores for the analysis methods employed.

Exploratory analyses were also carried out to examine whether intervention effects held for Hispanic children alone, and whether the effect of instruction depended on the Hispanic children's initial English fluency. Subsequent analyses examined whether there were differences in intervention effects relative to students' grade levels at the beginning of the study. Grade was not included in the initial analyses because the number of participants in some cells in the three-way analysis would have been quite small. We then examined whether gender or basis for recruitment into the study (ag-

gressive vs. low reading score) interacted with condition in their effects on the five reading measures.

The Effect of Instruction and Its Interaction with Ethnicity

Table 3 presents the means and standard deviations of the difference scores on the three reading measures obtained at Time 1 and Time 4 and means and standard deviations for the two measures that were obtained at Time 4 but not at Time 1. Table 4 presents the results of the two-way analysis involving intervention condition and ethnicity for each measure. Table 4 also presents Cohen's d , the standardized mean difference for each of the subtests (Cohen, 1988; Rosenthal & Rosnow, 1991). Effect size reports the importance or magnitude of the impact of a treatment, and it is relevant for research that seeks to evaluate the effectiveness of particular interventions.

The effect of instruction on letter-word identification was not significant, $F(1, 190) = 2.14$, ns , $d = .22$, although children who received supplemental instruction had greater increases in letter-word identification. There was a significant effect for instruction on word attack, with treatment children performing significantly better than their matched controls, $F(1, 188) = 10.30$, $p < .01$, $d = .46$, and we found a significant interaction between condition and ethnicity, $F(1, 188) = 4.80$, $p < .05$, $d = .32$. Analysis of simple main effects and independent analysis of Hispanics and non-Hispanics revealed no significant improvement for non-Hispanics, $F(1, 71) = .06$, ns , $d = .06$, but strong improvement among Hispanics, $F(1, 117) = 16.38$, $p < .001$, $d = .74$. We found a significant effect on oral reading fluency, with children who received supplemental instruction making greater gains on the number of words read per minute, $F(1, 189) = 7.46$, $p < .01$, $d = .40$. Treatment groups did not differ significantly on vocabulary, $F(1, 190) = 2.81$, ns , $d = .24$, or comprehension, $F(1, 188) = 3.19$, ns , $d = .26$, although the effects approached significance ($p < .10$) for both. Except for word attack, we found no significant interactions between treatment condition and ethnicity.

Effects for Hispanic Students Alone

Because the effectiveness of this intervention for Hispanic children has significant policy implications, we also examined intervention effects for the Hispanic children alone. These analyses distinguished Hispanic children who could speak English at the outset of the study from those who could not and included this as a second independent variable in the analyses. Table 5 presents those results. We found no effect of instruction on letter-word identification or vocabulary, but supplemental instruction improved student performance on measures of word attack, $F(1, 115) = 16.55$, $p < .001$, $d = .76$, oral reading fluency, $F(1, 115) = 6.01$, $p < .05$, $d = .46$, and passage comprehension, $F(1, 114) = 4.13$, $p < .05$, $d = .38$. In no case was there a significant interaction between Hispanic

TABLE 3. Means and Standard Deviations for Gain Scores T1–T4, and T4 Vocabulary and Comprehension Scores

Dependent variable			Ethnicity		
			Non-Hispanic	Hispanic	All participants
Woodcock-Johnson Letter-Word Identification gain score (T1-T4)					
Control	<i>M</i>	14.86	21.15	18.85	
	<i>SD</i>	18.51	18.78	18.83	
	<i>n</i>	35	61	96	
Intervention	<i>M</i>	20.54	25.05	23.26	
	<i>SD</i>	21.87	26.03	24.44	
	<i>n</i>	39	59	98	
Woodcock-Johnson Word Attack gain score (T1-T4)					
Control	<i>M</i>	5.24	1.41	2.78	
	<i>SD</i>	21.86	19.74	20.49	
	<i>n</i>	34	61	95	
Intervention	<i>M</i>	6.51	16.97	12.76	
	<i>SD</i>	24.45	22.17	23.56	
	<i>n</i>	39	58	97	
Oral Reading Fluency gain score (T1-T4)					
Control	<i>M</i>	47.43	46.54	46.87	
	<i>SD</i>	28.27	27.52	27.66	
	<i>n</i>	35	60	95	
Intervention	<i>M</i>	57.44	60.32	59.17	
	<i>SD</i>	35.78	33.50	34.27	
	<i>n</i>	39	59	98	
Woodcock-Johnson Vocabulary T4					
Control	<i>M</i>	31.51	20.38	24.44	
	<i>SD</i>	21.44	15.89	18.78	
	<i>n</i>	35	61	96	
Intervention	<i>M</i>	37.92	24.28	29.66	
	<i>SD</i>	25.86	19.45	23.06	
	<i>n</i>	39	60	99	
Woodcock-Johnson Comprehension T4					
Control	<i>M</i>	36.60	25.39	29.56	
	<i>SD</i>	25.47	15.38	20.36	
	<i>n</i>	35	59	94	
Intervention	<i>M</i>	39.77	32.08	35.11	
	<i>SD</i>	24.87	19.67	22.07	
	<i>n</i>	39	60	99	

Note. All scores are reported as Normal Curve Equivalent (NCE) scores.

TABLE 4. Analyses of Variance for Reading Scores by Condition and Ethnicity

Dependent variable	F (p values) Effect Size, d			Residual	
	Condition (C)	Ethnicity (E)	C × E	df	MS
Woodcock-Johnson Letter-Word Identification	2.139 (.145) .22	2.782 (.097) .24	.076 (.783) .04	190	475.161
Woodcock-Johnson Word Attack	10.301 (.002) .46	1.189 (.277) .16	4.808 (.030) .32	188	478.120
Woodcock-Johnson Vocabulary	2.813 (.095) .24	12.717 (.000) .52	.230 (.632) .06	190	392.141
Woodcock-Johnson Comprehension	3.191 (.076) .26	6.599 (.011) .38	.297 (.586) .08	188	423.520
Oral reading fluency	7.462 (.007) .40	.052 (.820) .04	.164 (.686) .06	189	982.065

Note. ANCOVAs were conducted on vocabulary and comprehension, using oral reading fluency at Time 1 as the covariate. All scores are reported as normal curve equivalent (NCE) scores.

TABLE 5. Analyses of Variance for Reading Scores by Condition and Hispanic Spanish/English Readers

Dependent variable	F (p values) Effect Size, d			Residual	
	Condition (C)	Spanish/ English (L)	C × L	df	MS
Woodcock-Johnson Letter-Word Identification	.925 (.338) .18	1.262 (.264) .20	.321 (.572) .10	116	514.0 79
Woodcock-Johnson Word Attack	16.548 (.000) .76	1.431 (.234) .22	.861 (.355) .18	115	438.1 92
Woodcock-Johnson Vocabulary	1.306 (.255) .22	1.776 (.185) .24	.755 (.387) .16	116	302.9 33
Woodcock-Johnson Comprehension	4.126 (.045) .38	.001 (.978) .00	.253 (.616) .10	114	318.6 24
Oral Reading Fluency	6.005 (.016) .46	2.103 (.150) .28	.208 (.649) .08	115	935.8 85

Note. ANCOVAs were conducted on vocabulary and comprehension, using oral reading fluency at Time 1 as the covariate.

children's initial English proficiency and the effects of instruction. Nevertheless, depending on the measure, we obtained scores for only 16 or 17 non-English-speaking children, so this study does not provide adequate power to detect those interactions. And, for four of the five variables the difference between conditions appeared larger for the non-English-speaking children than for those who spoke English at the outset. The findings of no interaction effects should be interpreted with caution, given the possibility of a Type II error.

Effects According to Grade

We conducted two-way ANOVAs and ANCOVAs with condition and initial grade level. The interaction effects in these analyses tested whether the effects of the intervention differed according to students' grade levels at the beginning of the study. For letter-word identification, $F(1, 186) = .634$, *ns*, $d = .20$; word attack, $F(1, 184) = .741$, *ns*, $d = .16$; oral reading fluency, $F(1, 185) = 2.26$, *ns*, $d = .38$; vocabulary, $F(1, 186) = .647$, *ns*, $d = .18$; and comprehension, $F(1, 184) = .515$, *ns*, $d = .22$, there were no significant interactions, suggesting that the effectiveness of the intervention was not mediated by initial grade. With oral reading fluency, however, there was an interaction that approached significance, $F(1, 185) = 2.26$, $p = .083$, $d = .38$, indicating that a more powerful test of this intervention (i.e., a study with more participants) may identify differences in the effectiveness of the intervention as a function of grade.

Discussion

The results support the utility of supplemental reading instruction for children in K and Grades 1, 2, and 3, and provide evidence of the effectiveness of Corrective Reading and Reading Mastery in improving the reading achievement of children at risk for reading difficulty. Hispanic and non-Hispanic children benefitted comparably from instruction, as measured by a significant effect on oral reading fluency, with an effect size of .40 and no interaction. The analysis of word attack indicated that Hispanic children primarily benefitted from the intervention. Although the main effects of treatment were not significant for vocabulary or comprehension, they were very nearly so ($p = .095$ and $p = .076$, respectively). It is possible that the outcome would have been favorable had the vocabulary and comprehension subtests been administered at pretest along with the other measures. The pretest measures would have allowed for a more accurate adjustment for random variation among participants before intervention.

Additional analyses examined Hispanic children alone. Regardless of students' initial English language facility, supplemental instruction enhanced word attack and the higher level literacy skills of oral reading fluency and passage comprehension. The corresponding effect sizes of .76, .46, and .38

indicated that Hispanic intervention students made significant growth relative to the Hispanic students in the control condition. What is more, the lack of facility with the English language was not detrimental to their responses to instruction. The findings suggest that the child's dominant language at the time of instruction was not a factor in their ability to benefit from instruction. These results are not consistent with contentions in the literature that English reading instruction is harmful for children for whom English is a second language (e.g., Crawford, 1989). On the contrary, they provide direct evidence that such instruction helps Spanish-speaking children develop reading skills that are known to be vital to their academic success (Slavin et al., 1996; Snow et al., 1998).

The direct instruction format, teacher modeling, and immediate feedback to students that characterize Reading Mastery and Corrective Reading seem to be of value to students at risk for reading disability. In particular, these curricula may be especially beneficial to English-language learners. The structured teaching, clear expectations, and progress monitoring may well provide the scaffolding that supports the development of reading skills in a second language. The results also attest to the important role of instructional assistants. With the limited time available to teachers to work with individual students, well-trained instructional assistants can make a valuable contribution to preventing or remediating reading difficulties.

Although the results indicate the benefit of 2 years of supplemental instruction, scores on the *Woodcock* subtests at Time 4 indicate that treatment children were, on average, still performing below their grade-level peers. This finding is consistent with an intervention study that investigated the effects of individual tutoring for early elementary children experiencing difficulties acquiring word-level reading skills (Torgesen, Wagner, & Rashotte, 1997). It was not feasible in the present study to give students the recommended 40 to 50 daily minutes of instruction (Engelmann & Bruner, 1988; Engelmann et al., 1988), and this could explain the aforementioned outcome. In any case, future investigations of supplemental instruction will likely need to provide more intensive and possibly longer term interventions to bring at-risk readers up to grade level.

In conclusion, the results of this 1-year follow-up are consistent with results obtained at the end of instruction (Gunn et al., 2000) in showing that supplemental instruction can prevent reading failure among Hispanic and non-Hispanic children. Benefits of supplemental instruction in reading were detectable a year after supplemental instruction had been discontinued. Effects on oral reading fluency were significant for both Hispanic and non-Hispanic children, and nearly significant effects on vocabulary and comprehension were also observed for both ethnic groups. For word attack skills and reading comprehension, there was clear benefit of instruction for Hispanic children. Gains did not depend on grade at which instruction began. Moreover, instructional benefits were equally apparent for Hispanic children regardless of whether or not

they spoke English at the outset of instruction. This finding should be viewed with caution, however, due to the small number of children who initially spoke only Spanish. This study provided only minimal power to detect interactions between language spoken and treatment condition among Hispanic students. In sum, it would appear that supplemental instruction using a direct instruction program is a viable, effective approach to preventing reading failure among at-risk children. The fact that most children were still performing below grade level at follow-up assessment, however, suggests that continued instruction may be required.

At a time when there is a pressing need to curb the disproportionately high dropout rate among older Hispanic students, it is critical to ensure that younger Hispanic students have the foundational skill they need to succeed in school—the ability to read. Students who read can learn from new materials and help ensure their own school success, even in the absence of quality instruction. Given the disparity in efficacy among instructional approaches used to teach reading, and the likelihood that this trend will continue, supplemental instruction is a viable option for helping struggling beginning readers.

AUTHORS' NOTES

1. Preparation of this manuscript was supported by Grant No. DA09678 from the National Institute on Drug Abuse.
2. The authors wish to thank Jason Blair and Claudia Montañó for coordinating data collection, and Wanda Ballentine for her assistance in the preparation of this manuscript.

REFERENCES

- Achenbach, T. M. (1991). *Manual for the teacher's report form and 1991 profile*. Burlington: University of Vermont, Department of Psychiatry.
- Adams, G. L., & Engelmann, S. (1996). *Research on direct instruction: 25 years beyond DISTAR*. Seattle, WA: Educational Achievement Systems.
- Adams, M. J. (1990). *Beginning to read: Thinking and learning about print*. Cambridge, MA: MIT Press.
- Anderson, R. C., Hiebert, E. H., Scott, J. A., & Wilkinson, I. A. G. (1985). *Becoming a nation of readers: The report of the Commission on Reading*. Washington, DC: National Institute of Education, U.S. Department of Education.
- Anderson, S., Auquier, A., Hauck, W., Oakes, D., Vandaele, W., & Weisberg, H. I. (1980). *Statistical methods for comparative studies. Techniques for bias reduction*. New York: Wiley.
- Baker, S. K., & Good, R. H., III. (1994, April). *Curriculum-based measurement reading with bilingual Hispanic students: A validation study with second-grade students*. Paper presented at the annual meeting of the Council for Exceptional Children/National Training Program for Gifted Education, Denver, CO.
- Biglan, A., Hood, D., Brozovsky, P., Ochs, L. M., Ary, D. V., & Black, C. (1991). Subject attrition in prevention research. In C. G. Leukefeld & W. J. Bukoski (Eds.), *Drug abuse prevention intervention research: Methodological issues NIDA research monograph 107* (pp. 213–233). Rockville, MD: National Institute on Drug Abuse.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.
- Crawford, J. (1989). *Bilingual education: History, politics, theory, and practice*. Trenton, NJ: Crane.
- Cunningham, A. E., & Stanovich, K. E. (1998). The impact of print exposure on word recognition. In J. L. Metsala & L. C. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 235–262). Mahwah, NJ: Erlbaum.
- Deno, S. L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children*, 52, 219–232.
- Ehri, L. C. (1998). Grapheme–phoneme knowledge is essential for learning to read words in English. In J. L. Metsala & L. C. Ehri (Eds.), *Word recognition in beginning reading* (pp. 3–40). Hillsdale, NJ: Erlbaum.
- Engelmann, S., & Bruner, E. C. (1988). *Reading mastery*. Chicago: Science Research Associates.
- Engelmann, S., Carnine, L., & Johnson, G. (1988). *Corrective reading: Word-attack basics—Teacher presentation book, Decoding A*. Chicago: Science Research Associates.
- Fitzgerald, J. (1995). English-as-a-Second Language reading instruction in the United States: A research review. *Journal of Reading Behavior*, 27, 115–152.
- Francis, D. J., Shaywitz, S. E., Stuebing, K. K., Shaywitz, B. A., & Fletcher, J. M. (1996). Developmental lag versus deficit models of reading disability: A longitudinal, individual growth curves analysis. *Journal of Educational Psychology*, 88, 3–17.
- Good, R. H., III, & Kaminski, R. A. (1996). Assessment for instructional decisions: Toward a proactive/prevention model of decision-making for early literacy skills. *School Psychology Quarterly*, 11, 326–336.
- Good, R. H., III, Kaminski, R., Laimon, D., & Johnson, D. S. (1992). Advances in assessment for the primary prevention of early academic problems. *The Oregon Conference Monograph, 1992* (pp. 100–109). Eugene: University of Oregon College of Education.
- Greenwood Institute. (1994). Notes on the Summary Report of the 1994 Washington Summit on Learning Disabilities. Retrieved November 24, 1999, from <http://www.greenwoodinstitute.org/resources/restat.html>
- 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. *The Journal of Special Education*, 34, 90–103.
- Hintze, J. M., Owen, S. V., Shapiro, E. S., & Daly, E. J. (2000). Generalizability of oral reading fluency measures: Application of G theory to curriculum-based measurement. *School Psychology Quarterly*, 15(1), 52–68.
- Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology*, 80, 437–447.
- Kameenui, E. J., Simmons, D. C., Baker, S. K., Chard, D. J., Dickson, S. V., Gunn, B. K., Smith, S., Sprick, M., & Lin, S. (1998). Effective strategies for teaching beginning reading. In E. J. Kameenui & D. W. Carnine (Eds.), *Effective strategies that accommodate diverse learners* (pp. 45–70). Columbus, OH: Merrill.
- Markell, M. A., & Deno, S. L. (1997). Effects of increasing oral reading: Generalization across reading tasks. *The Journal of Special Education*, 31, 233–250.
- Marston, D. (1989). Curriculum-based measurement: What is it and why do it? In M. Shinn (Ed.), *Curriculum-based measurement: Assessing special children* (pp. 18–78). New York: Guilford.
- Moats, L. C. (1995). The missing foundation in teacher education. *American Educator*, 19(2), 9–51.
- President's Advisory Commission on Educational Excellence for Hispanic Americans. (1996). Call to action: An executive summary. www.ed.gov/pubs/FaultLine/call.html. U.S. Department of Education.
- Rosenthal, R., & Rosnow, R. L. (1991). *Essentials of behavioral research, methods and data analysis*. San Francisco: McGraw-Hill.
- Slavin, R. E., Madden, N. A., Dolan, L. J., Wasik, B. A., Ross, S., Smith, L., & Dianda, M. (1996). Success for all: A summary of research. *Journal of Education for Students Placed at Risk*, 1(1), 41–76.
- Snow, C. E., Burns, S. M., & Griffin, P. (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy of Sciences, National Research Council, Commission on Behavioral and Social Sciences and Education.

- Stahl, S., & Miller, P. (1989). Whole language and language experience approaches for beginning reading: A quantitative research synthesis. *Review of Educational Research*, 59, 116.
- Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, 21, 360-406.
- Stotsky, S. (2000). The state of the state standards in English. In C. E. Finn & P. J. Petrilli (Eds.), *The state of the state standards*. Dayton, OH: Fordham Foundation.
- Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (1997). Approaches to the prevention and remediation of phonologically based reading disabilities. In B. Blachman (Ed.), *Foundations of reading acquisition and dyslexia: Implications for early intervention* (pp. 287-304). Mahwah, NJ: Erlbaum.
- U.S. Department of Education & U.S. Department of Justice. (1998). The potential of after-school programs. *Safe and smart: Making after-school hours work for kids*. <http://www.ed.gov/pubs/SafeandSmart/chapter1.html>. Washington, DC: U.S. Department of Education.
- Wagner, R. K., Torgesen, J. K., Laughon, P., Simmons, K., & Rashotte, C. A. (1993). Development of young readers' phonological processing abilities. *Journal of Educational Psychology*, 85, 83-103.
- Wagner, R. K., Torgesen, J. K., Rashotte, C. A., Hecht, S. A., Barker, T. A., Burgess, S. R., Donahue, J., & Garon, T. (1997). Changing causal relations between phonological processing abilities and word-level reading as children develop from beginning to fluent readers: A five-year longitudinal study. *Developmental Psychology*, 33, 468-479.
- Walker, H., & McConnell, S. (1988). *The Walker-McConnell scale of social competence and school adjustment*. Austin: PRO-ED.
- Walker, H., & McConnell, S. (1993). *Technical manual for the Walker-McConnell Scale of Social Competence and School Adjustment*. Austin: PRO-ED.
- Wilson, M. J., & Bullock, L. M. (1989). Psychometric characteristics of behavior rating scales: Definitions, problems, and solutions. *Behavioral Disorders*, 14, 186-200.
- Wolf, M., & Bowers, P. G. (1999). The double-deficit hypothesis for the developmental dyslexias. *Journal of Educational Psychology*, 91, 415-438.
- Woodcock, R. W., & Johnson, M. B. (1989). *Woodcock-Johnson psycho-educational battery-Revised*. Allen, TX: DLM Teaching Resources.
- Woodcock, R. W., & Mather, N. (1989, 1990). WJ-R tests of achievement: Examiner's manual. In R. W. Woodcock & M. B. Johnson (Eds.), *Woodcock-Johnson psycho-educational battery-Revised*. Allen, TX: DLM Teaching Resources.

