Effective Reading Programs for English Language Learners and Other Language-Minority Students

Alan Cheung
Success for All Foundation
Center for Data-Driven Reform in Education

Robert E. Slavin
Johns Hopkins University
Center for Data-Driven Reform in Education

Abstract

This article systematically reviews research on elementary reading programs for English language learners and other language-minority students. It focuses on studies that compared experimental and control groups on quantitative reading measures. Among beginning reading models, research supported structured, phonetic programs emphasizing language development in both native-language and English instruction. Tutoring programs were also supported. For upper-elementary reading, research supported a broad range of programs, but particularly effective were programs using cooperative learning, extensive vocabulary instruction, and literature.

Introduction

For many years, the focus of policy debates relating to the reading education of English language learners (ELLs) has been on the question of language of instruction, contrasting bilingual and English-only approaches. As important as language of instruction is, however, there has been a growing recognition in recent years that *quality* of instruction is at least as important as *language* of instruction in the ultimate success of ELLs (see, for example, August & Hakuta, 1997; Brisk, 1998; Christian & Genesee, 2001; Goldenberg, 1996; Secada, Chavez-Chavez, García, Munoz, Oakes, Santiago-Santiago, & Slavin, 1998).

Research on language of instruction, reviewed most recently by Greene (1997) and Slavin and Cheung (in press), has generally found that bilingual programs are more effective than English-only programs. Slavin and Cheung found particularly strong evidence favoring paired bilingual programs, in which students are taught to read both in their native language and in English, beginning in kindergarten or first grade, a strategy typically seen in two-way bilingual programs. However, in today's political environment, the language of reading instruction is likely to be determined by factors beyond the control of individual educators. Whatever the language of instruction may be, educators concerned with ELLs need programs known to be effective with these students.

Quality of instruction is the product of many factors, including the quality of teachers, class size, and other resources. One factor is the program of instruction used each day to teach reading. A number of coherent, replicable reading programs combining materials and professional development have been developed and used with ELLs. This article reviews research on reading programs for ELLs and other language-minority students in an attempt to apply consistent, well-justified standards of evidence to draw conclusions about which of these programs are effective for these children. The review applies a technique called "best-evidence synthesis" (Slavin, 1986), which seeks to apply consistent, clear standards to identify unbiased, meaningful information from experimental studies, and then discusses each qualifying study, computing effect sizes but also describing the context, design, and findings of each study. Best-evidence synthesis closely resembles meta-analysis, but it requires more extensive discussion of key studies. Details of this procedure are described later. The purpose of this review is to examine the quantitative evidence on replicable reading programs for ELLs and other language-minority students to discover how much of a scientific basis there is for competing claims about effects of various programs. Our purpose is to inform practitioners, policymakers, and researchers about the current state of the evidence on this topic as well as gaps in the knowledge base in need of further scientific investigation.

Review Methods and Criteria for Inclusion

Review methods for studies of reading programs for ELLs and other language-minority students were as follows:

1. The studies involved elementary (K–6) children identified as ELLs or language-minority (e.g., "Hispanic") students in English-speaking countries.
2. The studies compared children taught in classes using a given reading program to those in control classes using standard textbooks.

3. The language of instruction was the same in experimental and control groups.
4. Random assignment or matching with appropriate adjustments for any pretest differences had to be used. Studies without control groups, such as pre–post comparisons and comparisons to "expected" gains, were excluded, as were studies with pretest differences of more than one standard deviation.
5. The dependent measures included quantitative measures of reading performance, such as standardized reading measures. In all cases, measures included assessments of comprehension, not just phonics or decoding. The focus on quantitative measures was intended to allow for comparable, objective conclusions about program effects across studies.
6. A minimum treatment duration of 12 weeks was required.

Studies of Beginning Reading Programs

It is in the earliest years of formal education that children define themselves as learners, largely on the basis of reading success. The early elementary years are of particular importance for ELLs, as this is the time when they are most likely to be struggling both to learn a new language and to learn to read. Perhaps because of this, the largest number of methodologically adequate studies have focused on the early elementary grades. Studies in this section are ones in which the treatments begin in kindergarten or first grade.

There were 13 studies of beginning reading that met the criteria outlined above. Most studies of reading approaches for ELLs and other language-minority students lacked control groups or objective measures, did not document or control for pretest differences, or were very brief. The main characteristics and findings of the qualifying studies are summarized in Table 1.

Success for All

Among the beginning reading studies that met the inclusion criteria, six evaluated the Success for All program (Slavin & Madden, 1999, 2001). Success for All is a comprehensive reform model that provides schools with well-structured curriculum materials emphasizing systematic phonics in Grades K–1, and cooperative learning, direct instruction in comprehension skills, and other elements in Grades 2–6. It also provides extensive professional development and follow-up for teachers, frequent assessment and regrouping, one-to-one tutoring for children who are struggling in reading, and family support programs. A full-time facilitator helps all teachers implement the model.

For ELLs, Success for All has two variations. One is a Spanish bilingual program, *Éxito Para Todos*, which teaches reading in Spanish in Grades 1–2 and then transitions them to English-only instruction, usually starting in third

Table 1

Beginning Reading Programs: Descriptive Information and Effect Sizes for Qualifying Studies

Study	Intervention description	Design	Duration	*N*	Grade
Success For All					
Nunnery, Slavin, Ross, Smith, Hunter, & Stubbs (1997)	Success for All (SFA)-Bilingual	Matched control	1 year	298 in 30 schools	1
Livingston & Flaherty (1997)	SFA-Bilingual	Matched control	3 years	6 schools (3 E & 3 C)	1–3
	SFA-English Language Development Adaptation				
	SFA-English Language Development Adaptation				
Slavin & Madden (1995)	SFA-English Language Development Adaptation	Matched control	5 years	50 in 2 schools	K
Slavin & Yampolsky (1991)					
Slavin, Leighton, & Yampolsky (1990)					
Ross, Smith, & Nunnery (1998)	SFA-English Language Development Adaptation	Matched control	1 year	540 in 6 schools	1
Hurley, Chamberlain, Slavin, & Madden (2001)	SFA	Compared gains to the state mean for Hispanic students	4 years	95 SFA schools	(K–2)—>(3–5)
Chambers, Slavin, Madden, Cheung, & Gifford (2004)	SFA with embedded video	Matched control	1 year	455 in 8 schools	K–1

Sample characteristics	Evidence of initial equality	Posttest	Effect size	Median effect size
Success For All				
Spanish-dominant students across 30 schools with bilingual programs in Houston, Texas	Fairly well matched on demographic and well matched on pretest. control group (C) > experimental group (E); effect size (ES) = -0.08	**Spanish Woodcock**		
		Word Identification	+0.24	+0.22
		Word Attack	+0.26	
		Passage Comprehension	+0.20	
Spanish-dominant bilingual students in California	Well matched on demographics and Peabody Picture Vocabulary Test (PPVT) pretests.	**Spanish Woodcock**		
		Grade 1		+0.97
		Grade 2		+0.44
		Grade 3		+0.03
Spanish-dominant ESL students in California	Well matched on demographics and PPVT pretests.	**English Woodcock**		
		Grade 1		+1.36
		Grade 2		+0.46
		Grade 3		-0.09
Other ESL students in California	Well matched on demographics and PPVT pretests.	**English Woodcock**		
		Grade 1		+0.24
		Grade 2		+0.37
		Grade 3		+0.05
Asian students in Philadelphia	Well matched on overall achievement level, poverty, and other variables.	**English Woodcock**	**Grade 4**	
		Word Identification	+1.54	+1.49
		Word Attack	+1.49	
		Passage Comprehension	+0.62	
		English Woodcock	**Grade 5**	
		Word Identification	+1.40	+1.33
		Word Attack	+1.33	
		Passage Comprehension	+0.75	
Tucson, Arizona: 39% Hispanic, 67% free lunch	Well matched on demographics and pretests.	**English Woodcock**		
		Word Identification	+0.51	+0.52
		Word Attack	+0.83	
		Passage Comprehension	+0.41	
		Durrell	+0.32	
Hispanic students in Texas	Well matched on initial Texas Assessment of Academic Skills (TAAS) reading scores.	**English TAAS Reading (Grades 3–5)**	+0.28*	+0.28* (ES from school means, not individual scores)
Hispanic students in New York City, Washington, DC, rural Arizona, and southern California	Well matched on PPVT.	**English Woodcock**		
		Word Identification	+0.40	+0.36
		Word Attack	+0.36	
		Passage Comprehension	+0.21	

Table 1, cont.

Beginning Reading Programs: Descriptive Information and Effect Sizes for Qualifying Studies

Study	Intervention description	Design	Duration	*N*	Grade
Other programs					
Chambers, Cheung, Madden, Slavin, & Gifford (2004)	Embedded video (SFA with embedded video vs. SFA)	Random assignment of schools	1 year	172 in 10 schools	1
Becker & Gersten (1982)	Direct Instruction	Matched control	Follow-up study—2 years after the treatment	225	K–3
Gersten (1985)	Direct Instruction	Matched control	8 months	~35	1–2
Stuart (1995)	Phonetic program (Jolly Phonics [JP]) vs Literature-based program (Big Books [BB])	Matched control	12 weeks	112	K
Escamilla (1994)	Reading Recovery in Spanish *(Descubriendo la Lectura)*	Matched control	7 months	46	1
Gunn, Biglan, Smolkowski, & Ary (2000)	Small group tutoring using Direct Instruction	Random assignment	2 years	122	K–4
Goldenberg (1990)	Use of teacher-created booklets at home and at school	Quasi-experimental	8 months	56	K

Sample characteristics	Evidence of initial equality	Posttest	Effect size	Median effect size
Other programs				
Hispanic students in Hartford, Connecticut	Well matched on PPVT, Word ID	**English Woodcock**		
		Word Identification	+0.23	+0.20
		Word Attack	+0.36	
		Passage Comprehension	+0.16	
		DIBELS[a] Fluency	+0.07	
Hispanic English language learner (ELL) students in Uvalde, Texas	Well matched on demographics	**English Wide Range Achievement Test (WRAT) Reading**	**Across 2 grades**	
		Level II	+0.44	+0.21
		Level I	+0.50	
		Mean	**+0.47**	
		English Metropolitan Achievement Test (MAT)		
		Word Knowledge	+0.11	
		Reading	+0.21	
		Total Reading	+0.16	
		Mean	**+0.16**	
Asian ELL students	Similar on LAS scores for cohort 1 (C>E) and cohort 2 (C>E)	**English Comprehensive Test of Basic Skills (CTBS) Reading**		
		Experimental	**75%**	E>C
		Control	**19%**	
		English CTBS Language		
		Experimental	**71%**	E>C
		Control	**44%**	
Sylheti-dominant students in London	Well matched on demographics but not on pretests; JP>BB; ES=+0.88 on phonics knowledge pretests; JP>BB; ES=+0.70 on reading and writing pretests	**English Woodcock**		
		Phoneme awareness (5 measures)	+0.70	Immediate tests: +0.88
		Delayed tests (1 year later)	+0.16	
		Reading and Spelling (5 measures)	+1.06	Delayed tests: +0.34
		Delayed tests (1 year later)	+0.52	
Spanish-dominant bilingual students in Arizona	Well matched on Spanish Aprenda, but on Spanish observation survey, C>E, median ES=-0.43 across four measures	**Spanish Woodcock**		
		Spanish *Aprenda*	+0.30	+0.30
		Spanish Observation Survey (6 measures)	+0.84	+0.84
Low-achieving Hispanic students in rural Oregon	Well-matched on English Woodcock-Johnson Letter Word Identification and Word Attack scales, and Oral Reading Fluency	**English Woodcock**		
		Year 1 Letter Word Word Attack Oral Reading Fluency	 +0.22 +0.10 +0.16	Year 1 +0.22
		Year 2 Letter Word Word Attack Oral Reading Vocabulary Comprehension	 +.046 +0.91 +0.43 +0.44 +0.48	Year 2 +0.44
Spanish-dominant students in southern California	Similar on Bilingual Syntax Measure and free lunch	**Spanish Woodcock**		
		13 measures of early literacy development	+0.83	+0.83

[a]DIBELS stands for the Dynamic Indicator of Basic Early Literacy Skills.

grade. The other is an English language development (ELD) adaptation, which teaches children in English with appropriate supports, such as vocabulary development strategies linked to the words introduced in children's reading texts.

Studies of Success for All with ELLs and other language-minority students have compared children taught using the Spanish adaptation to other children taught in Spanish, or have compared the ELD adaptation to other ELD English reading programs.

Success for All: Spanish Bilingual Adaptation (*Éxito Para Todos*)

California (bilingual)

Researchers at the Southwest Educational Research Laboratory (now part of WestEd) conducted a 3-year longitudinal study involving three California elementary schools and three matched controls. They pooled data across the schools in four categories: English-dominant students, Spanish-dominant students taught in Spanish, Spanish-dominant students taught in English, and speakers of languages other than English or Spanish taught in English. Three cohorts were followed. Data for a 1992 cohort were reported for Grades 1, 2, and 3; for 1993, Grades 1 and 2; and for 1994, Grade 1 only.

Students in the two *Éxito Para Todos* schools in California scored higher on the Spanish Woodcock than controls at every grade level in all three cohorts (Livingston & Flaherty, 1997; Dianda & Flaherty, 1995). Median effect sizes across cohorts averaged +0.97 for first graders, +0.44 for second graders, and +0.03 for third graders. The analyses for second and third graders understate the magnitude of the differences. In line with district and program policies, students initially taught in Spanish were transitioned into English instruction as soon as they demonstrated an ability to succeed in English. Because of their success in Spanish reading, many more *Éxito Para Todos* than control students were transitioned during second and third grades. Therefore, the highest achieving experimental students were being removed from the Spanish sample, reducing the mean for this group. (This is a common problem in studies of transitional bilingual education.)

Houston (bilingual)

The largest study of *Éxito Para Todos* took place in the Houston Independent School District in Texas. Both Spanish and English forms of Success for All were studied (see Nunnery, Slavin, Ross, Smith, Hunter, & Stubbs, 1997).

The bilingual portion of the study compared first graders in 20 schools implementing *Éxito Para Todos* to those in 10 matched schools also using Spanish bilingual instruction. Children were assessed on three scales from the Spanish Woodcock: Word Identification, Word Attack, and Passage

Comprehension. Ten children were selected at random from each school; after missing data were removed, there were 298 Spanish-dominant students across the 30 schools with bilingual programs.

School-level comparisons showed significant differences ($p < .05$) between Success for All schools and comparison schools on Word Identification and Word Attack. Overall, the median student-level effect size (ES) in comparison to controls was +0.22.

Success for All: English Language Development Adaptation

Philadelphia (English language development)

The first evaluation of the ELD adaptation of Success for All took place at Francis Scott Key Elementary in Philadelphia (Slavin & Madden, 1995). Sixty-two percent of Key's students were from Asian backgrounds, primarily Cambodian. Nearly all of them entered the school in kindergarten with little or no English. The remainder of the school was divided between African American and White students.

The program was evaluated in comparison to a matched Philadelphia elementary school. The two schools were very similar in overall achievement level and other variables. All students in Grades 4–5, most of whom had been in their respective programs since kindergarten, were individually administered three scales from the Woodcock Language Proficiency Battery (Woodcock, 1984): Word Identification, Word Attack, and Passage Comprehension. Asian Success for All students at both grade levels performed substantially better than Asian control students. The differences were statistically significant on every measure at both grade levels ($p < .001$).

California (English language development)

The 3-year California study (Livingston & Flaherty, 1997; Dianda & Flaherty, 1995) included data on ELLs taught in English. These included both students in one Modesto school that did not have a bilingual program, as well as ELLs in the two schools (one in Modesto and one in Riverside) who were speakers of languages other than English or Spanish.

Results for Spanish-dominant students taught in English showed strong impacts for first graders (ES = +1.36), smaller ones for second graders (ES = +0.46), and no differences for third graders (ES = -0.09). Again, the transitioning of successful students out of English as a Second Language (ESL) classes reduced the apparent differences by third grade (because the highest achieving students were no longer receiving ESL services).

Results for speakers of languages other than English or Spanish (taught in English) were similar to those for Spanish-dominant ESL students. Averaging across cohorts, effect sizes were +0.24 for first graders, +0.37 for second graders, and +0.05 for third graders (Livingston & Flaherty, 1997; Dianda & Flaherty, 1995).

Arizona (English language development)

Another study of the ELD adaptation of Success for All in schools serving many students acquiring English took place in an Arizona school district (Ross, Smith, & Nunnery, 1998). This 1-year study compared first graders in two Success for All schools to those in four schools using locally developed Title I schoolwide projects. Students were pretested on the English Peabody Picture Vocabulary Test (PPVT) and then posttested on the Woodcock Word Identification, Word Attack, and Passage Comprehension scales, and the Durrell Oral Reading Test. Analyses of covariance found that Hispanic Success for All students scored significantly higher than control students on all measures (ES = +0.52).

Texas Statewide Evaluation

Hurley, Chamberlain, Slavin, and Madden (2001) reported an analysis of data from the Texas Assessment of Academic Skills (TAAS), comparing reading gains (from the year schools began to implement Success for All to 1998) by all 111 Success for All schools in the state to those made by students throughout Texas. The comparisons involving Hispanic students are relevant to this review. Note that while the TAAS data were for Grades 3–5, most of the students had been in the program 3 to 4 years, meaning that they had begun in Grades K–2.

Ninety-five of the Success for All schools had enough Hispanic students in Grades 3–5 to be included in the analysis. Analyzing at the school level, their TAAS reading gains were significantly greater ($p < .01$) than those for Hispanic students in the state as a whole. Hispanic students in the SFA schools and state means for Hispanic students were similar in the year before SFA was introduced. The effect size for school means was +0.28.

Success for All With Embedded Video

Chambers, Slavin, Madden, Cheung, and Gifford (2004) carried out a study of an adaptation of Success for All that incorporated embedded video. Four types of video material were used: animations to present letter sounds, puppet vignettes to present sound blending, live-action skits to present vocabulary, and a variety of segments from the television program *Between the Lions* to reinforce various skills. The brief video segments were interspersed in teacher's lessons in Grades K–1. Hispanic students were expected to benefit in particular from the SFA and embedded video treatment because the videos included vocabulary presentations and clear visual reinforcements of reading skills. Hispanic students in four schools in different parts of the United States were compared to matched students in similar schools that did not use Success for All or embedded video. A yearlong study involving 311 experimental and 144 control students in Grades K–1 found that, controlling for PPVT, schools

using Success for All with embedded video scored significantly higher than controls on Woodcock Word Identification (ES = +0.40), Word Attack (ES = +0.36), and Passage Comprehension (ES = +0.21).

Success for All: Conclusions

The effects of Success for All on the achievement of ELLs and other language-minority students are not entirely consistent, but in general they are substantially positive. Across two studies of *Éxito Para Todos*, the Spanish bilingual adaptation of Success for All, the median effect sizes on Spanish assessments was +0.41. Across five studies of the ELD adaptation of Success for All, the median effect size was +0.37.

Embedded Video

A recent study compared Success for All schools using the embedded video materials described above to schools also implementing Success for All but without the embedded videos (Chambers, Cheung, Madden, Slavin, & Gifford, 2004). Because all schools used SFA, this was not a study of Success for All but of the added embedded video treatment. Ten majority-Hispanic schools in inner-city Hartford, Connecticut, were randomly assigned to SFA plus embedded video or SFA-only (control) conditions for a 1-year experiment. Results for Hispanic children, who were 66% of the sample, found positive effects controlling for the PPVT and the Woodcock Word Identification scale on Woodcock Word Identification (ES = +0.23) and Word Attack (ES = +0.36).

Direct Instruction

Direct Instruction (DI), or DISTAR (Adams & Engelmann, 1996), is a reading program that starts in kindergarten with very specific instructions to teachers on how to teach beginning reading skills. It uses reading materials with a phonetically controlled vocabulary, rapidly paced instruction, regular assessment, and systematic approaches to language development. Like Success for All, DI provides extensive professional development and coaching to all teachers. DI was not specifically written for ELLs or Latino students, but it is often used with them.

The most important evaluation of DI was the Follow Through study of the 1970s, in which nine early literacy programs were evaluated (Stebbins, St. Pierre, Proper, Anderson, & Cerva, 1977). In sites throughout the United States, matched experimental and control schools were compared on various measures of reading.

One of the sites was in Uvalde, Texas, which primarily served Hispanic students. Becker & Gersten (1982) carried out a follow-up of the Follow Through study when the children who had experienced the treatments in Grades K–3

were in Grades 5–6. They found that the Uvalde DI students, who were well matched on demographic factors with their control group, scored substantially better than the controls. Effect sizes averaged +0.47 for two scales of the individually administered Wide Range Achievement Test (WRAT) and +0.16 across three Metropolitan Achievement Test (MAT) subscales, for a median across five tests in two grades of ES = +0.21.

Gersten (1985) evaluated DI as part of a structured immersion program for limited English proficient students who spoke a variety of Asian languages. In addition to the DI beginning reading program, the structured immersion model emphasized English at a level understood by the students, occasional translation, preteaching of vocabulary, and direct teaching of the structure of the English language. Students in a matched control group participated in programs whose characteristics were not described, but which also primarily taught in English.

Across two cohorts, 75% of DI students scored at or above grade level on the Comprehensive Test of Basic Skills (CTBS) Total Reading Scale at the end of 2 years, while only 19% of comparison students were at or above grade level ($p < .001$).

Jolly Phonics (Systematic Phonics Instruction)

Stuart (1999) carried out an evaluation of Jolly Phonics (JP), an English phonetic kindergarten reading program, in five London primary schools. This program was compared to a big books program emphasizing teaching by drawing children's attention to letters and words in popular children's stories. The subjects were mostly ELLs, and among these, most were speakers of Sylheti (a language of Bangladesh). Most subjects were 5-year-olds. One teacher in each school volunteered to implement either JP or Big Books (BB). The JP and BB schools were well matched on most variables, including free meals and academic performance, but the JP schools had many more children at beginning ESL levels (53% vs. 30%).

The interventions took place 1 hour per day for 12 weeks. The results strongly favored the JP group. Effect sizes for five gain scores measures of phonemic awareness and phonics knowledge had a median value of +0.70 at posttest and +0.16 on a delayed posttest 1 year later. On five measures of reading and writing, the median effect size for gain scores was +1.06 at the end of the experiment and +0.52 1 year later.

Reading Recovery/*Descubriendo la Lectura*

Reading Recovery is an early intervention tutoring program for young readers who are experiencing difficulty in their first year of reading instruction (Clay, 1993). The program provides the lowest achieving readers (lowest 20%) in first grade with supplemental tutoring in addition to their regular reading classes. Children participating in Reading Recovery receive daily one-to-one

30-minute lessons for 12–20 weeks with a certified, specially trained teacher. The lessons include assessment, reading known stories, reading a story that was read once the day before, writing a story, working with a cut-up sentence, and reading a new book. *Descubriendo la Lectura* (DLL), the Spanish adaptation of Reading Recovery, is equivalent in all major aspects to the original program. There have been many evaluations comparing Reading Recovery and control students, including a large-scale, randomized evaluation in Ohio (Pinnell, Lyons, Deford, Bryk, & Seltzer, 1994). Only one study involving ELLs met the inclusion standards of this review. This was a 7-month evaluation of DLL conducted by Escamilla (1994) in Tucson, Arizona. The experiment compared 23 DLL students to 23 matched comparison students also taught in Spanish in another school. In both cases, students were identified as being in the lowest 20% of their classes based on individually administered tests and teacher judgment. The two groups were well matched on the Spanish *Aprenda*. The outcomes of DLL on Spanish reading measures at the end of first grade were very positive. On six scales of a Spanish Observation Survey adapted from the measures used in evaluations of the English Reading Recovery program, DLL students started out below controls and ended the year substantially ahead of them, with a median effect size of +0.84.

Small Group Tutorials with Direct Instruction

Gunn, Biglan, Smolkowski, and Ary (2000) evaluated a small-group tutorial program that used two forms of DI, *Reading Mastery* and *Corrective Reading*, as a supplementary intervention for Hispanic and non-Hispanic children who were struggling in reading. The children were in kindergarten to third grade, and were selected either because they scored at a very low level on an achievement measure or because they were rated by their teachers as being high in aggressive behavior (and were below grade level in reading). Children were selected from nine rural Oregon elementary schools. They were randomly assigned to experimental or control conditions. Those children assigned to the experimental group were taught in homogeneous groups of one to three children using *Reading Mastery* if they were in Grades K–2, or *Corrective Reading* if they were in Grades 3–4. They were taught daily by instructional assistants for 2 years. Only 19 of the 122 Hispanic students were considered non-English speaking; the oral English skills of the remaining students were not specified.

The experimental and control groups were very well matched on the Woodcock-Johnson Letter Word Identification and Word Attack scales, and on Oral Reading Fluency. After the first year, tutorial students who had received 5 to 6 months of supplementary instruction showed greater gains than control students on all three measures, Letter–Word ID (ES = +0.22), Word Attack (ES = +0.70), and Fluency (ES = +0.16). Only the Word Attack differences were significant. At the end of the second year, after 15–16 months of instruction,

effect sizes for gains from pretest on these measures were +0.46, +0.91, and +0.43, respectively. In addition, there were positive effects on Woodcock Reading Vocabulary (ES = +0.44) and Passage Comprehension (ES = +0.48), given as posttests only. Experimental–control differences on all five measures were significant after 2 years.

Libros

Goldenberg (1990) studied a school and home reading intervention for Spanish-dominant kindergartners. The intervention, called *Libros*, involved teachers introducing and extensively discussing a Spanish story and then sending home photocopied "books" with children once every 3 weeks through kindergarten. Parents were encouraged to read with their children and were shown a videotape of a parent reading and discussing the story. In control classrooms, teachers sent home worksheets on letters and syllables. Children in four classrooms using *Libros* were matched with those in four control classrooms based on Bilingual Syntax Measure scores. On an experimenter-constructed set of 13 Spanish early literacy assessments at the end of the year, experimental children scored significantly higher than controls (median ES = +0.83). Effects were strongest on measures of letter and word identification but were less positive on comprehension measures.

Studies of Upper Elementary Reading Programs

Several studies have evaluated reading programs for ELLs in Grades 2–5. Seven of these met the inclusion criteria. These are summarized in Table 2 and described in the following sections.

Bilingual Cooperative Integrated Reading and Composition

An experiment by Calderón, Hertz-Lazarowitz, and Slavin (1998) evaluated a cooperative learning program called Bilingual Cooperative Integrated Reading and Composition, or BCIRC. BCIRC is an adaptation of Cooperative Integrated Reading and Composition (CIRC), an upper elementary reading program based on principles of cooperative learning, that has been successfully evaluated in several studies (see Stevens, Madden, Slavin, & Farnish, 1987). BCIRC was adapted to meet the needs of limited English proficient children in bilingual programs who are transitioning from Spanish to English reading. In CIRC and BCIRC, students work in four-member heterogeneous teams. After a teacher introduction, students engage in a set of activities related to a story they are reading. These include partner reading in pairs, and team activities focused on vocabulary, story grammar, summarization, reading comprehension, creative writing, and language arts. BCIRC adds to these activities transitional readers (in this study, Macmillan's *Campanitas de Oro* and *Transitional Reading*

Program) and ESL strategies, such as total physical response, realia, and appropriate use of cognates, to help children transfer skills from Spanish to English reading.

Control teachers also used the same *Campanitas de Oro* and *Transitional Reading Program* textbooks, and received training in generic cooperative learning strategies. None of the control teachers used cooperative learning consistently, although all of them made occasional use of these strategies.

The BCIRC study involved 222 Hispanic children in the Ysleta Independent School District in El Paso, Texas. Seven of the highest poverty schools in the district were assigned to experimental (three schools) or control (four schools) conditions. As a whole, the experimental and control schools were well matched demographically. Two cohorts were assessed, one of which was involved for just 1 year (second grade) and the other for 2 years (Grades 2–3). Analyses of covariance controlling for Bilingual Syntax Measure scores found significantly higher scores for students in BCIRC classes in both cohorts.

Bilingual Transition With Success for All

An experiment by Calderón, August, Slavin, Durán, Madden, and Cheung (2004) evaluated an enriched transition program for children who had been taught in Spanish using Success for All and were moving to the English program in third grade. The enriched program, a descendent of BCIRC, included an English phonics program called FastTrack Phonics, rapidly presented components of the Success for All beginning reading (Reading Roots) program including the embedded videos described earlier, and explicit instruction in vocabulary using strategies similar to those used by Carlo, August, McLaughlin, Snow, Dressler, Lippman, Lively, et al. (2004). The experiment compared students in El Paso, Texas, who received the full program to matched students in similar control schools. After 1 year, students in the program scored higher than control students (controlling for Spanish and English Woodcock Scales) on Woodcock Word Attack (ES = +0.21), Passage Comprehension (ES = +0.16), and Picture Vocabulary (ES = +0.11). Experimental students scored higher on some of the Spanish measures as well.

Enriched Transition

Saunders and Goldenberg (1996) evaluated a program designed to help ELLs transition from Spanish to English. The treatment focused on literature study, writing, discourse, skill building, reading comprehension strategies, independent reading, teacher read-alouds, and other elements. These treatments were applied to second and fifth graders in transitional bilingual education (TBE) and English-only classes. In each case, a control group was matched with the experimental group. Over a year, the English-only

Table 2

Upper Elementary Reading Programs: Descriptive Information and Effect Sizes for Qualifying Studies

Study	Intervention description	Design	Duration	*N*	Grade	Sample characteristics
Calderón, Hertz-Lazarowitz, & Slavin (1998)	Bilingual Cooperative Integrated Reading & Composition (BCIRC)	Matched control	2 years	222	2–3	Spanish-dominant students in El Paso, Texas
Calderón, August, Slavin, Durán, Madden, & Cheung (2004)	Success for All with enriched transition	Matched control	1 year	239 in 8 schools	3	Spanish-dominant students in El Paso, Texas
Saunders & Goldenberg (1996)	Enriched transition	Matched control	1 year	140	2 & 5	Spanish-dominant students in southern California
Saunders & Goldenberg (1999)	Enriched transition	Matched control	3 years	102	1–5	Spanish and Cantonese speaking students in southern California

Evidence of initial equality	Posttest	Effect size	Median effect size
Well matched on demographics and pretests	**Spanish Texas Assessment of Academic Skills (TAAS)**	**Grade 2**	Spanish Reading +0.30 Spanish Writing +0.62 English Reading +0.54 English Writing/ Language +0.29
	Reading	+0.30	
	Writing	+0.62	
	English TAAS	**Grade 3**	
	Reading	+0.54	
	Writing	+0.29	
	English TAAS	**2 years**	
	Reading	+0.87	
	Language	+0.38	
	English TAAS	**1 year**	
	Reading	+0.33	
	Language	+0.22	
Well matched on English and Spanish Woodcock measures			
	English Woodcock		
	Picture Vocabulary	+0.11	+0.16
	Passage Comprehension	+0.16	
	Word Attack	+0.21	
Well matched on pretests	**English-only group**		
	2nd grade-English Reading	+0.34	+0.19
	5th grade-English Reading	+0.03	
	Transitional bilingual education group		
	2nd grade-Spanish Reading	+1.36	+1.36
	5th grade-English Reading	+0.68	+0.68
Well matched on % of limited English proficient, socioeconomic status, ethnicity, and achievement scores	**Spanish subgroup**	**Spanish measures**	
		Reading	Language
	1st grade	-0.02	+0.11
	2nd grade	+0.26	+0.20
	3rd grade	+0.38	+0.27
	4th grade	+0.59	+0.38
	Cantonese subgroup	**English measures**	
	4th grade	+0.53	+1.77
	5th grade	+0.80	+0.78

Table 2, cont.,
Upper Elementary Reading Programs: Descriptive Information and Effect Sizes for Qualifying Studies

Study	Intervention description	Design	Duration	*N*	Grade	Sample characteristics
Carlo et al. (2004)	Direct instruction in key vocabulary	Matched control	2 years	~130	4–5	English language learner (ELL) students in California, Virginia, and Massachusetts
Pérez (1981)	Oral-language activity	Matched control	3 months	150	3	Mexican American ELL students in Texas
Denton, Anthony, Parker, & Harsbrouck (2004)	Read Well (tutoring using systematic phonics)	Random assignment	4 months	33	2–5	Spanish-dominant bilingual students in Texas
	Read Naturally (tutoring using repeated readings)			60		

Evidence of initial equality	Posttest	Effect size	Median effect size
Well matched on pretests	**English Vocabulary Assessment**		+0.21
	Peabody Picture Vocabulary Test (PPVT)	-0.08	
	Polysemy production	+0.33	
	Morphology	+0.22	
	Semantic Association	+0.21	
	English Reading Comprehension	+0.17	
Well matched on demographics and pretests, experimental group (E) > control group (C), effect size (ES) = +0.15	**English**	+0.97	+0.97
Well matched on Woodcock Reading Mastery Test (WRMT) pretests; E > C, ES = +0.32 (0.3 < p < 0.6)	**English—Read Well**		+0.51
	Word Identification	+0.55	
	Word Attack	+0.46	
	Passage Comprehension	+0.00	
	Fluency	+0.18	
	Accuracy	+0.78	
	Comprehension	+0.82	
Well matched on WRMT pretests	**English—Read Naturally**		+0.08
	Word Identification	-0.05	
	Word Attack	-0.13	
	Passage Comprehension	+0.16	
	Fluency	+0.23	
	Accuracy	+0.30	
	Comprehension	+0.00	

experimental group scored higher than control groups on an English reading measure in second grade (ES = +0.34) but not in fifth grade (ES = +0.03). Second-grade TBE students, tested in Spanish, scored substantially better in the experimental condition (ES = +1.36). Fifth-grade experimental TBE students tested in English also showed substantially higher achievement (ES = +0.68).

The Saunders and Goldenberg (1996) article only reported on the first year of a 3-year transition project. A study of the full program was described by Saunders (1998). It compared children in the 3-year transition program (using the methods described above) to those in a 3- to 6-month transition, the usual treatment for ELLs in the district studied. On Spanish measures, differences were insignificant in Grade 1 (ES = -0.02) and Grade 2 (ES = +0.26), but significant in Grade 3 (ES = +0.38) and Grade 4 (ES = +0.59). In a Cantonese-dominant subgroup, experimental students scored significantly higher on English tests (Grade 4, ES = +0.53; Grade 5, ES = +0.80). At fifth grade, an early-transitioning group was tested in English and a late-transitioning group was tested in Spanish. In both cases, effects favored the experimental group (ES = +0.50 for English, ES = +0.92 for Spanish). Similar effects were seen on performance measures of reading and writing, and experimental students passed a test used as a criterion for placement in English-only instruction at much higher rates than did controls.

Vocabulary Intervention

Carlo et al. (2004) carried out a 2-year evaluation of a vocabulary teaching intervention with Spanish-dominant fourth and fifth graders in California, Massachusetts, and Virginia. The intervention involved introducing 12 vocabulary words each week, using a variety of strategies, such as charades, 20 questions, discussions of Spanish cognates, word webs, and word association games.

The experimental students were taught in one 5-week unit and two 6-week units in the first year, and three 5-week units in the second year. Matched control students continued their usual instruction. Experimental and control students were not significantly different on any of an extensive set of measures.

At the end of the first year, ELLs showed greater gains from pretest than controls, but surprisingly, gains were lower after 2 years of intervention.

Pérez (1981) evaluated an oral-language intervention with Spanish-dominant third graders in Texas. The intervention consisted of daily 20-minute sessions in which children worked with humorous language games, pictures, and other activities intended to build their oral language. The experimental group of 75 students was compared to a well-matched control group. On an unspecified reading measure, the experimental group scored substantially higher.

Tutoring

Two types of one-to-one tutoring for ELLs were studied in an experiment by Denton, Anthony, Parker, and Hasbrouck (2004). Spanish-dominant students in Grades 2–5 in a bilingual program in Texas were assigned to one of two separate experiments. Those scoring lower than the first-grade level on the Woodcock Word Attack scale were randomly assigned to a program called Read Well (Sprick, Howard, & Fidanque, 1998), or to an untutored control group. Those scoring higher than this were randomly assigned to a tutoring program called Read Naturally or to an untutored control group. Read Well uses systematic phonics instruction and practice in fully decodable text (like the first-grade instruction in Success for All). Read Naturally (Ihnot, 1992) emphasized repeated readings of connected text, vocabulary, and comprehension instruction. Tutors were undergraduate education majors. All tutoring was done in English. The final sample of students in the Read Well evaluation included 19 experimental and 14 control children. Those in the experimental group received an average of 22 tutoring sessions. In the Read Naturally comparison, there were 32 tutored and 28 non-tutored children.

The results indicated substantially higher achievement for the Read Well students than for controls, with a median effect size of +0.51 across six measures. Differences were statistically significant only on the Woodcock Word Attack scale ($p < .05$) and an oral reading accuracy scale ($p < .001$). In contrast, there were no differences between the children tutored with Read Naturally and those who were not tutored (ES = +0.08).

Conclusions: Studies of Reading

The research summarized in this article shows how much remains to be done on effective reading programs for ELLs and other language-minority students. Only a handful of studies met the minimal inclusion standards applied in this review, which principally required an experimental-control comparison of a reading program over at least 12 weeks, with evidence that the two groups were equivalent at pretest.

Beginning Reading

Among the 13 studies of interventions beginning in kindergarten or first grade that met the inclusion standards, the evidence supported structured, phonetic programs emphasizing language development, in both native-language and English instruction. The largest number of studies involved Success for All, a comprehensive reform model (Slavin & Madden, 1999). Two studies of Success for All in its Spanish bilingual form found consistent positive effects on students' Spanish reading performance, with a median effect size of +0.41 (in comparison to schools teaching in Spanish using

alternative methods). Similarly, five studies of schools using the English-language adaptation of Success for All with Latino and Asian ELLs found positive effects, with a median effect size of +0.37.

Studies evaluating Success for All with embedded video materials found positive effects of the combined program for Hispanic students (Chambers, Slavin, et al., 2004) and found that the embedded videos added significantly to the effects (Chambers, Cheung, et al., 2004).

Two longitudinal studies found strong and lasting effects of DI on the reading achievement of language-minority students. One was a follow-up of mostly Hispanic fifth and sixth graders in Texas who had experienced DI in Grades K–3 (Becker & Gersten, 1982). The other was a 2-year study of DI in a structured immersion program for Asian ELLs (Gersten, 1985). An adaptation of DI for use in small-group tutorials (one to three children) also found positive effects (Gunn et al., 2000).

No other beginning reading program had more than a single methodologically adequate study. A study of the systematic phonics program JP (Stuart, 1999) found promising effects among children of Bangladeshi origin in London, but the study had serious problems with pretest differences. Very positive effects were documented in a study of a Spanish adaptation of Reading Recovery (Escamilla, 1994). A study of *Libros*, a home and school literature approach using Spanish reading materials, documented benefits for ELL kindergartners (Goldenberg, 1990).

Upper Elementary Reading

Seven studies of reading in Grades 2–5 met the inclusion criteria. The evidence generally supported programs that make extensive use of cooperative learning, vocabulary instruction, and literature. A 2-year evaluation of BCIRC (Calderón et al., 1998), a cooperative learning strategy, found strong positive effects on the Spanish and English reading of children transitioning from Spanish to English reading in Grades 2–3. A similar treatment, an enriched Spanish-to-English transition program based on Success for All, also showed significantly positive effects on English reading performance (Calderón et al., 2004). Saunders (1998), and Saunders and Goldenberg (1999), successfully evaluated an enriched transition process for ELLs moving to English-only instruction. Carlo et al. (2004) found positive effects of an English vocabulary intervention for ELL fourth and fifth graders on various experimenter-made measures of vocabulary skill, and Pérez (1981) found that instruction in oral English skills improved the reading skills of ELL third graders. Denton et al. (2004) evaluated two tutoring approaches and found that Read Well, a phonetic program, improved the English reading of very low–achieving ELLs.

It is important to note that the programs with the strongest evidence of effectiveness in this review are all programs that have also been found to be effective with students in general: Success for All (Slavin & Madden, 2000,

2001), DI (Adams & Engelmann, 1996), Reading Recovery (Pinnell et al., 1994), and phonetic tutoring (e.g., Wasik & Slavin, 1993). In fact, several of the studies evaluating Success for All (e.g., Nunnery et al., 1997; Livingston & Flaherty, 1997; Ross et al., 1998), as well as DI (Gunn et al., 2000), also included non-ELL students, and in each case those students also gained from the interventions, to about the same degree. The beginning reading programs with the strongest evidence of effectiveness in this review made use of systematic phonics, such as Success for All, DI, and JP, but systematic phonics has been identified as a component of effective beginning reading programs for English-proficient students as well (see National Reading Panel, 2000; Gersten & Geva, 2003). Typically, programs originally designed for use with English-proficient students are considerably adapted for use with ELLs, with more emphasis on vocabulary and oral language (see Fitzgerald, 1995; Slavin & Calderón, 2001).

While we do have a good start on research in several areas, there is much more to be done. Large-scale, randomized, longitudinal evaluations of well-justified approaches are needed to more confidently recommend effective strategies for ELLs and other language-minority students of all ages and backgrounds. Research systematically varying program components and research combining quantitative and qualitative methods are needed to more fully understand how various interventions affect the development of reading skills among ELLs. It is time to end the ideological debates, and to instead focus on good science, good practice, and sensible policies for children whose success in school means so much to themselves, their families, and our nation's future.

References

Adams, G. L., & Engelmann, S. (1996). *Research on Direct Instruction: 25 years beyond DISTAR.* Seattle, WA: Educational Achievement Systems.

August, D., & Hakuta, K. (1997). *Improving schooling for language-minority children: A research agenda.* Washington, DC: National Research Council.

Becker, W. C., & Gersten, R. (1982). A follow-up on Follow Through: The later effects of the Direct Instruction model on children in fifth and sixth grades. *American Educational Research Journal, 19*(1), 75–92.

Brisk, M. E. (1998). *Bilingual education.* Mahwah, NJ: Erlbaum.

Calderón, M., August, D., Slavin, R. E., Durán, D., Madden, N. A., & Cheung, A. (2004). *The evaluation of a bilingual transition program for Success for All.* Baltimore: Johns Hopkins University, Center for Research on the Education of Students Placed at Risk.

Calderón, M., Hertz-Lazarowitz, R., & Slavin, R. E. (1998). Effects of Bilingual Cooperative Integrated Reading and Composition on students making the transition from Spanish to English reading. *Elementary School Journal, 99*(2), 153–165.

Carlo, M. S., August, D., McLaughlin, B., Snow, C. E., Dressler, C., Lippman, D., et al. (2004). Closing the gap: Addressing the vocabulary needs of English language learners in bilingual and mainstream class-rooms. *Reading Research Quarterly, 39*(2), 188–215.

Chambers, B., Cheung, A., Madden, N. A., Slavin, R. E., & Gifford, R. (2004). *Achievement effects of embedded multimedia in a Success for All reading program.* Baltimore: Johns Hopkins University, Center for Research on the Education of Students Placed at Risk.

Chambers, B., Slavin, R. E., Madden, N. A., Cheung, A., & Gifford, R. (2004). *Effects of Success for All with embedded video on the beginning reading achievement of Hispanic children.* Baltimore: Johns Hopkins University, Center for Research on the Education of Students Placed at Risk.

Christian, D., & Genesee, F. (Eds.). (2001). *Bilingual education.* Alexandria, VA: TESOL.

Clay, M. M. (1993). *Reading Recovery: A guidebook for teachers in training*. Portsmouth, NH: Heinemann.

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

Dianda, M., & Flaherty, J. (1995, April). *Effects of Success for All on the reading achievement of first graders in California bilingual programs*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Escamilla, K. (1994). Descubriendo la Lectura: An early intervention literacy program in Spanish. *Literacy, Teaching, and Learning, 1*(1), 57–70.

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

Gersten, R. (1985). Structured immersion for language minority students: Results of a longitudinal evaluation. *Educational Evaluation and Policy Analysis, 7*(3), 187–196.

Gersten, R., & Geva, E. (2003). Teaching reading to early language learners. *Educational Leadership, 60*(8), 44–49.

Goldenberg, C. (1990, April). *Evaluation of a balanced approach to literacy instruction for Spanish-speaking kindergartners*. Paper presented at the annual meeting of the American Educational Research Association, Boston.

Goldenberg, C. (1996). The education of language-minority students: Where we are, and where do we need to go? *Elementary School Journal, 36*(4), 715–738.

Greene, J. P. (1997). A meta-analysis of the Rossell & Baker review of bilingual education research. *Bilingual Research Journal, 21*(2&3), 103–122.

Gunn, B., Biglan, A., Smolkowski, K., & Ary, D. (2000). The efficacy of supplemental instruction in decoding skills for Hispanic and non-Hispanic students in early elementary school. *The Journal of Special Education, 34*(2), 90–103.

Hurley, E. A., Chamberlain, A., Slavin, R. E., & Madden, N. A. (2001). Effects of Success for All on TAAS Reading: A Texas statewide evaluation. *Phi Delta Kappan, 82*(10), 750–756.

Ihnot, C. (1992). *Read Naturally*. St. Paul, MN: Read Naturally.

Livingston, M., & Flaherty, J. (1997). *Effects of Success for All on reading achievement in California schools*. Los Alamitos, CA: WestEd.

National Reading Panel. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. Rockville, MD: National Institute of Child Health and Human Development.

Nunnery, J., Slavin, R., Ross, S., Smith, L., Hunter, P., & Stubbs, J. (1997, March). *Effects of full and partial implementations of Success for All on student reading achievement in English and Spanish.* Paper presented at the annual meeting of the American Educational Research Association, Chicago.

Pérez, E. (1981, October). Oral language competence improves reading skills of Mexican-American third graders. *Reading Teacher, 35*(1), 24–27.

Pinnell, G., Lyons, C. A., DeFord, D. E., Bryk, A. S., & Seltzer, M. (1994). Comparing instructional models for the literacy education of high risk first graders. *Reading Research Quarterly, 29*, 9–40.

Ross, S. M., Smith, L. J., & Nunnery, J. A. (1998, April). *The relationship of program implementation quality and student achievement.* Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.

Saunders, W. M. (1998). *Improving literacy achievement for English learners in transitional bilingual programs*. Long Beach, CA: Center for Research on Education, Diversity, and Excellence, University of California.

Saunders, W. M., & Goldenberg, C. (1996). *The effects of a comprehensive language arts transition program on the literacy development of English language learners*. Santa Cruz: Center for Research on Education, Diversity, and Excellence, University of California.

Saunders, W. M., & Goldenberg, C. (1999). *The effects of a comprehensive language arts transition program on the literacy development of English language learners*. Santa Cruz: Center for Research on Education, Diversity, and Excellence, University of California.

Secada, W. G., Chavez-Chavez, R., García, E., Munoz, C., Oakes, J., Santiago-Santiago, I., et al. (1998). *No more excuses: The final report of the Hispanic dropout project*. Washington, DC: U.S. Department of Education.

Slavin, R. E. (1986). Best-evidence synthesis: An alternative to meta-analytic and traditional reviews. *Educational Researcher, 15*(9), 5–11.

Slavin, R. E., & Calderón, M. (Eds.) (2001). *Effective programs for Latino students*. Mahwah, NJ: Erlbaum.

Slavin, R. E., & Cheung, A. (in press). A synthesis of research on language of reading instruction for English language learners. *Review of Educational Research*.

Slavin, R. E., & Madden, N. A. (1995). *Effects of Success for All on the achievement of English language learners*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans.

Slavin, R. E., & Madden, N. A. (1999). Effects of bilingual and English as a second language adaptations of Success for All on the reading achievement of students acquiring English. *Journal of Education for Students Placed at Risk, 4*(4), 393–416.

Slavin, R. E., & Madden, N. A. (2000). Research on achievement outcomes of Success for All: A summary and response to critics. *Phi Delta Kappan, 82*(1), 38–40, 59–66.

Slavin, R. E., & Madden, N. A. (2001). *One million children: Success for All*. Thousand Oaks, CA: Corwin.

Sprick, M. M., Howard, L. M., & Fidanque, A. (1998). *Read Well: Critical foundations in primary reading*. Longmont, CO: Sopris West.

Stebbins, L. B., St. Pierre, R. G., Proper, E. C., Anderson, R. B., and Cerva, T. (1977). *Education as experimentation: A planned variation model, Vol. IV–A. An evaluation of Follow Through*. Cambridge, MA: Abt Associates.

Stevens, R. J., Madden, N. A., Slavin, R. E., & Farnish, A. M. (1987). Cooperative Integrated Reading and Composition: Two field experiments. *Reading Research Quarterly, 22*, 433–454.

Stuart, M. (1999). Getting ready for reading: Early phoneme awareness and phonics training improves reading and spelling in inner-city second language learners. *British Journal of Educational Psychology, 69*(4), 587–605.

Wasik, B. A., & Slavin, R. E., (1993). Preventing early reading failure with one-to-one tutoring: A review of five programs. *Reading Research Quarterly, 28,* 178–200.

Woodcock, R. W. (1984). *Woodcock Language Proficiency Battery*. Allen, TX: DLM.

Acknowledgments

We would like to thank Margarita Calderón, Nancy Madden, Bette Chambers, Diane August, and Tim Shanahan for their comments on an earlier draft.

This paper was written under funding from the Institute of Education Sciences, U.S. Department of Education (Grant No. R–305A–040082). However, any opinions expressed are those of the authors and do not necessarily reflect Department of Education positions or policies.