

Effective Reading Programs for Spanish Dominant English Language Learners (ELLs) in the Elementary Grades: A Synthesis of Research

Alan C. K. Cheung

Johns Hopkins University
200 W Towsontown Blvd
Towson, MD 21204
Tel: 410-616-2410
Email: acheung@jhu.edu

Robert E. Slavin

Johns Hopkins University
and University of York
200 W Towsontown Blvd
Towson, MD 21204
Tel: 410-616-2310
Email: rslavin@jhu.edu

March, 2012

Abstract

This review synthesizes research on English reading outcomes of all types of programs for Spanish-dominant ELLs in elementary schools. It is divided into two major sections. One focuses on studies of language of instruction, and one on reading approaches for ELLs other than bilingual education. A total of 14 qualifying studies met the inclusion criteria for language of instruction. Though the overall findings indicate a positive but modest effect ($ES=+0.19$) in favor of bilingual education, the largest and longest-term evaluations, including the only multiyear randomized evaluation of transition bilingual education, did not find any differences in outcomes by the end of elementary school for children who were either taught in Spanish and transitioned to English or taught only in English. The review also identified some proven and promising whole-school and whole-class interventions, including Success for All, cooperative learning, Direct Instruction, and ELLA. In addition, programs that use phonetic small group or one-to-one tutoring have also shown positive effects for struggling readers. What is in common across the most promising interventions is their use of extensive professional development, coaching, and cooperative learning. The findings support a conclusion increasingly being made by researchers and policy makers concerned with optimal outcomes for ELLs and other language minority students: Quality of instruction is important than language of instruction.

Introduction

The number of English Language Learners (ELLs) has been increasing rapidly in the past few decades in the United States and will no doubt continue to rise. According to the National Clearinghouse for English Language Acquisition (2011), there were over 5 million ELLs in the United States in 2009, making up 10% of all K-12 students, compared to 3.5 million a decade ago. The percentage is expected to rise to 25% by 2030. The majority of ELLs nationally are from Spanish-language backgrounds. Based on the 2005 American Community Survey by the Modern Language Association, there were about 52 million speakers of languages other than English in the United States. Out of all non-native English speakers, Spanish speakers are by far the largest group (32 million or 62%). No other language is spoken by more than 3%.

In comparison to their non-ELL counterparts, ELLs tend to be at higher risk of performing poorly in early literacy. As their oral English improves, so does their English reading, but many ELLs are not able to catch up with their non-ELL counterparts as time progresses. On the most recent 2011 National Assessment of Educational Progress (2011), only 7% of fourth-grade ELLs scored at or above the proficient level, while 46% of non-ELLs scored this well. Among eighth graders, only 3% of ELLs scored at or above the proficient level, as compared to 39% of non-ELLs.

With the rate of immigration on the rise, teachers are facing enormous challenges in knowing how to best serve and educate ELLs in their schools. A critical and contentious issue in the education of ELLs is language of instruction. In the 1970s and 1980s, bilingual programs to teach ELLs had been common in many places. With the English-only movement in the late 1990s, several states passed propositions that enacted policies against the use of bilingual education, including California in 1998, Arizona in 2000, and Massachusetts in 2002. Though these propositions usually included waivers for parents who wanted their children to be in bilingual education, they were designed to make such waivers difficult. For example, after Proposition 227 was passed in California, the proportion of ELLs receiving primary language instruction with English language development dropped significantly, from 30% to 8%. To evaluate the effects of the implementation of Proposition 227 on ELLs, the California Department of Education contracted with the American Institutes for Research (AIR) and WestEd to carry out a non-experimental evaluation. No sizable effect of Proposition 227 was found on LEP students' academic achievement in English (Parrish et al., 2006). Similar results were also found in a study of *Question 2*, the Massachusetts English immersion law, on third-grade LEP students' reading achievement (Guo, in press).

With the passage of the *No Child Left Behind Act* (NCLB) in 2002, the use of bilingual education has been further discouraged throughout the U.S. For one thing, NCLB requires all states to include all ELLs in state testing programs that assess their academic skills in English, usually by third grade. Wright (2007) argued that “the high-stakes testing policies of NCLB, along with the accountability provisions which demand that “limited English proficient” students learn English as quickly as possible, ultimately serve to discourage schools from offering heritage language programs.”

The fundamental question has been whether ELLs should be taught using their native language or are better served in an English-only learning environment. Opponents of bilingual education argue that ELLs are better served by early and intensive exposure to an all-English learning environment (e.g., Rossell & Baker, 1996). On the other hand, bilingual advocates believe that ELLs are best served if they are gradually transitioned from their native language to English-only, because they can start with success in a language they understand and then what they learn in their native language can transfer as they learn English (Goldenberg, 1996; Thomas & Collier, 1997). A dozen reviews have been conducted on the relative effectiveness of bilingual education and structured immersion programs. Conclusions of these reviews have been quite diverse (Greene, 1997; Rossell & Baker, 1996; Slavin & Cheung, 2005; Willig, 1987). However, the recent evidence seems to suggest that the quality of instruction may be more important than the language of instruction (August & Hakuta, 1997; Christain & Genesee, 2001; Slavin, Madden, Calderon, Chamberlain, & Hennessy, 2011), and in any case, the focus of research and policy has shifted toward identifying effective strategies for helping ELLs succeed in English rather than just focusing on initial language of instruction.

The purpose of this review is to review effective reading interventions or strategies for Spanish-dominant ELLs, including native-language instruction as one among an array of means of potentially improving English reading. The overall focus on Spanish-dominant ELLs is justified by two factors. First, Spanish-speaking students are by far the largest minority ELL group in our public school systems. In addition, they have historically low educational attainment and a high dropout rate. According to 2009 data, the high school dropout rate for U.S. Hispanics was highest (17.6%) among all minority groups, as compared with African American (9.6%), White (5.2%), and Asian (2.1%) students (Child Trends Data Bank, 2011). Clearly, the U.S. cannot reach its national educational goals unless educators can greatly improve outcomes for this large and growing group.

Working definitions of ELLs and types of language of instruction

English language learners (ELLs). This term describes students who are in the process of acquiring English language skills and knowledge. Some schools refer to these students using the term limited-English-proficient (LEP), and English learners (ELs) is becoming common.

Language minority students. The term “language minority students” is used to refer to students whose parents speak a language other than English at home, but who may or may not have limited English proficiency themselves. This broader term is often used to define study populations when individual data on English proficiency are not available.

English-only programs. These programs focus mainly on English language development, and all instruction and activities are conducted in English. The goal is English language acquisition and academic achievement in English. A typical equivalent is structured English immersion (SEI), reflecting the idea that even when native language plays little or no role in reading instruction, ELLs are supported in their acquisition of English reading and speaking.

Transitional bilingual programs. These programs provide most instruction in students’ native language (L1) in the early grades, then gradually transition into an all-English (L2) learning environment in later grades.

Two-way bilingual immersion programs. These programs provide instruction in both L1 and the second language (L2) for ELLs and non-ELLs in the same classes. The goal is for both ELLs and native English-speaking students to become bilingual and biliterate (Genesee, 1999; Genesee, Paradis, & Crago, 2004).

Paired bilingual programs. These programs provide reading instruction to ELLs in both Spanish and English at different times of the day. They differ from two-way bilingual programs mainly in that English-proficient students are not necessarily taught in Spanish.

Previous Reviews on Language of Instruction

Several major meta-analyses of the impact of bilingual education on reading have been conducted in the past two decades (Greene, 1997; Rossell & Baker, 1996; Slavin & Cheung, 2005; Willig, 1987). The results were mixed. For example, Rossell & Baker examined 72 studies from the 1960s onward by using a vote-counting method. They concluded that most studies did not favor bilingual education. However, Greene (1997), Willig (1987), and Slavin & Cheung (2005) came to a different conclusion, concluding that bilingual programs produced better reading results for ELLs. For example, Greene used a meta-analysis method to examine the same studies that were included in Rossell & Baker (1996). He reported that only 11 out of the 72 studies included in the Rossell & Baker review were methodologically adequate. Greene found an overall effect size of +0.21 in support of bilingual programs among the methodologically adequate studies. Consistent with Greene's findings, Slavin & Cheung (2005) found a positive effect of bilingual programs, especially paired bilingual programs, on English reading achievement, with an overall effect size of +0.31. It is important to mention that few long-term randomized studies were included in these reviews. Also, most of these studies were done long ago, especially in the 1970's. In 2005, in an effort to produce a more satisfying answer to the long-standing debate on bilingual education, the U.S. Department of Education funded three large-scale longitudinal studies that used rigorous research designs to examine the relative effectiveness of transitional bilingual education (TBE) and structured English immersion (SEI). Results of these three longitudinal studies have appeared in the past few years (Francis & Vaughn, 2009; Irby et al., 2010; Slavin, et al., 2011). With this new evidence, there is a need to revisit the review of research on language of instruction.

Methods

The current review employed the best evidence synthesis review technique proposed by 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. Comprehensive Meta-analysis Software Version 2 (Borenstein, Hedges, Higgins, & Rothstein, 2005) was used to calculate effect sizes and to carry out various meta-analytical tests, such as Q statistics and sensitivity analyses. Like many previous research reviews, this study follows five key steps: 1. locating all possible studies; 2. screening potential studies for inclusion using preset criteria; 3. coding all qualified studies based on their methodological and substantive features; 4. calculating effect sizes for all qualified studies for further combined analyses; and

5. carrying out comprehensive statistical analyses covering both average effect sizes and the relationships between effect sizes and study features.

Literature Search Procedures

In an attempt to locate every study that could possibly meet the inclusion criteria, a literature search of articles written between 1970 and 2011 was carried out. Electronic searches were made of educational databases (e.g., JSTOR, ERIC, EBSCO, Psych INFO, Dissertation Abstracts), web-based repositories (e.g., Google Scholar), and ELL reading program providers' websites, using different combinations of key words. Descriptors included *bilingual education*, *structured immersion programs*, *English language learners*, *language of instruction*, *language minority students*, *English immersion*, *dual language*, *two-way bilingual education*, *English as a second language*, *effective reading program*, *reading intervention*, *elementary reading*, and *secondary reading*. We also conducted searches by program name. We attempted to contact producers and developers of ELL reading programs to check whether they knew of studies that we had missed. References from other reviews of language of instruction and effective reading programs for ELLs were further investigated. We also conducted searches of recent tables of contents of key journals from 2000 to 2011: *Reading Research Quarterly*, *American Educational Research Journal*, *Journal of Educational Research*, *Journal of Adolescent & Adult Literacy*, *Journal of Educational Psychology*, *Bilingual Research Journal*, and *Reading and Writing Quarterly*. Citations in the articles from these and other current sources were located.

Language of Instruction and Effective Reading Programs

The review is divided into two major sections. One focuses on studies of language of instruction (e.g., bilingual vs. English-only instruction), and one on reading approaches for ELLs other than bilingual education.

Inclusion Criteria for the Instructional Language Review

In order to be included in the review of language of instruction, studies had to meet the following inclusion criteria (see Slavin, 2008, for rationales).

1. The studies compared children taught reading in bilingual classes to those taught in English immersion classes, as defined earlier.

2. Either random assignment to conditions was used, or pretesting or other matching criteria established the degree of comparability of bilingual and immersion groups before the treatments began. If these matching variables were not identical at pretest, analyses adjusted for pretest differences or data permitting such adjustments were presented. Studies without control groups, such as pre-post comparisons or comparisons to “expected” scores or gains, were excluded. Studies with pretest differences exceeding half of a standard deviation were excluded.

A special category of studies was rejected based on the requirement of pretest measurement before treatments began. These were studies in which the bilingual and immersion programs were already under way before pretesting or matching. For example, Danoff, Coles, McLaughlin, & Reynolds (1978), in a widely cited study, compared one-year reading gains in many schools using bilingual or immersion methods. The treatments began in kindergarten or first grade, but the pretests (and later, posttests) were administered to children in grades 2-6. Because the bilingual children were primarily taught in their native language in K-1 and the immersion children were taught in English, their pretests in second grade would surely have been affected by their treatment condition. Meyer & Feinberg (1992, p.24) noted the same problem with reference to the widely cited Ramirez et al. study (1991), which also obtained pretests after students had been in bilingual or English-only programs: “It is like watching a baseball game beginning in the fifth inning: If you are not told the score from the previous innings, nothing you see can tell you who is winning the game.” Studies that tested children in upper elementary or secondary grades who had experienced bilingual or English-immersion programs in earlier years were included if premeasures were available from before the programs began, but in most cases such premeasures were not reported, so there is no way to know if the groups were equivalent beforehand (examples include Cuirel, Stenning, & Cooper-Stenning, 1980; Thomas & Collier, 2002).

3. The subjects were Spanish-dominant English language learners in elementary schools in the U.S. Studies that identified children as “language minority” (i.e., they came from homes in which Spanish was spoken but may or may not have been ELLs themselves) were included if data were not available on the language proficiency of individual children. Studies that mixed ELLs and English monolingual students in a way that did not allow for separate analyses were excluded (e.g., Skoczytas, 1972). Studies of children learning a foreign language were not included (e.g., monolingual English speakers studying Spanish). In addition, studies that involved languages other than Spanish were excluded (Morgan, 1971).

4. The dependent variables included quantitative measures of English reading performance, such as standardized tests and informal reading inventories. If treatment-specific measures were used, they were included only if there was evidence that all groups focused equally on the same outcomes. Measures of outcomes related to reading, such as language arts, writing, and spelling, were not included.
5. The treatment duration was at least one school year. For the reasons discussed later, even one-year studies of transitional bilingual education are less than ideal, because students taught in their native language are unlikely to have transitioned to English by the end of the study. Studies even shorter than this do not address the question in a meaningful way.

Both the first and second author looked at each potential study independently. When disagreements arose, both authors reexamined the studies in question together and came to a final agreement.

Effect Size Calculations and Statistical Analyses

In general, effect sizes were computed as the difference between experimental and control individual student posttests after adjustment for pretests and other covariates, divided by the unadjusted posttest pooled SD. Procedures described by Lipsey & Wilson (2001) and Sedlmeier & Gigerenzor (1989) were used to estimate effect sizes when unadjusted standard deviations were not available, as when the only standard deviation presented was already adjusted for covariates or when only gain score SD's were available. If pretest and posttest means and SD's were presented but adjusted means were not, effect sizes for pretests were subtracted from effect sizes for posttests. F ratios and t ratios were used to convert to effect sizes when means and standard deviations were not reported. After calculating individual effect sizes for all qualifying studies, Comprehensive Meta-Analysis software (Borenstein, et al., 2005) was used to carry out all statistical analyses such as Q statistics and overall effect sizes.

Findings: Language of Instruction

Overall findings

=====
Insert Table 1 here
=====

A total of 14 qualifying studies based on approximately 2,000 elementary school children met the inclusion criteria for language of instruction (see Table 1). The findings indicate a positive but modest effect ($ES=+0.19$), weighted by sample size, in favor of bilingual education. It is important to note that the majority of the included studies were conducted in the 1970s and most used a model of paired bilingual program that is quite different from those that are commonly used today. Unlike typical bilingual transition programs, students in these paired bilingual programs were taught reading in both English and Spanish at different times of the day.

Out of the 14 qualifying studies, there were only two long-term longitudinal studies¹ (Maldonado, 1977; Slavin et al., 2011). These two studies were of great importance because of their use of random assignment and their long durations. Randomized experiments avoid selection bias, a serious problem when parents or teachers decide whether children within the school are initially taught in Spanish or English. In addition, many studies comparing TBE and SEI were too brief to have given students in TBE sufficient time to make their transition to English. The Maldonado (1977) and Slavin et al. (2011) studies are described in detail below; for detailed description of the other included studies, please see Slavin and Cheung (2005).

The first 5-year longitudinal study was carried out by Maldonado (1977) with a group of Spanish-speaking Mexican-American elementary school children in Corpus Christie, Texas. The main objective of the study was to investigate how well the language-minority students were able to succeed in the regular education program of the school district after they had left the bilingual program. A total of 126 children in six elementary schools participated in the study. The treatment group was comprised of 47 children who had participated in the bilingual program for four consecutive years, from 1st grade to 4th

¹ Only part of the 4-year matched study of early-exit TBE carried out by Ramirez et al (1991) was included. The longitudinal aspect of the study has been excluded due to inadequate controls for pretest differences (Slavin & Cheung, 2005; Meyer & Fienberg, 1992).

grade. The control group consisted of 79 students enrolled in regular English-only classrooms for the same 4 years and grades. These two groups were followed until they reached fifth grade, one full year after the treatment group left the bilingual program. The treatment students received a minimum of 2 hours of instruction in Spanish daily in language arts, reading, mathematics, and social studies. However, the author did not provide specific information about the control condition. After controlling for pretest difference, no statistically significant differences were found between the treatment and control group at any grade. The final effect size for the fifth grade results was +0.11 (directionally favoring bilingual education). It is important to mention that teachers in both conditions were bilingual. However, it is not stated how much these bilingual teachers in the control condition used Spanish in their classrooms to help children who were in need of bilingual explanations. As the author stated, “It is highly possible that the control group bilingual teachers might have used the Spanish language for clarification of some concepts. This in turn would not only assist those students in the comprehension of those concepts but at the same time lower the difference between the groups in the areas of mathematics and reading” (p.104).

The second 5-year longitudinal study, conducted by Slavin and his colleagues (2010), was one of the three longitudinal studies funded by the U.S. Department of Education in 2005. The other two studies used 2x2 factorial designs to examine the effects of both language of instruction and an enhanced classroom intervention within each language of instruction (Francis, York, August, & Vaughn, 2009; Irby et al., 2010). In each case, the bilingual vs. English-only factor involved matching, not random assignment. We included the parts of the studies that examined the effectiveness of the enhanced intervention within each language of instruction, which will be discussed in the second part of this paper. However, the comparisons that assessed the relative effectiveness of TBE and SEI did not meet the inclusion criteria. There were large pretest differences ($ES > 1.00$) in oral language composite scores between the SEI and bilingual groups in the Francis et al. (2009) study, suggesting that children were more likely to be selected into SEI if their English was already good. In the Irby et al. (2010) study, only a select group of TBE students (25%) participated in the English TAKS posttesting as compared to all students in the SEI group.

Slavin et al. (2011) also compared the effectiveness of TBE and SEI. Six schools located in Los Angeles, California; Denver, Colorado; Albuquerque, New Mexico; St. Paul, Minnesota; Rockford, Illinois, and Alamo, Texas participated in the study. All participating schools had both transitional bilingual and structured English immersion programs. The study used a randomized within-school design in which kindergarteners were randomly assigned to either a transitional bilingual program or a structured

immersion program. To increase the sample size, three successive cohorts (children entering kindergarten in 2004, 2005, and 2006) were included and were pretested in the fall of kindergarten and then assessed each spring on both Spanish and English reading tests. Children in the TBE classes were initially taught reading in Spanish. Transition to English reading could begin as early as first grade, but the majority of the participating schools did not start the transition until second grade. By fourth grade, all children were taught in English entirely. Children in the SEI classes were taught in an English-only environment except for occasional Spanish explanations. Since children in both conditions used the Success for All reading program as their instructional materials, the content being taught was basically the same. The only difference between the two treatments was the language of instruction. The initial sample size was 247 (130 TBE and 117 SEI), and the final sample size was 115 (60 TBE; 55 SEI) due to attrition. No statistically significant difference was found between the two groups in terms of their attrition rate or the pretest scores of the final samples. Both the Peabody Picture Vocabulary Test and its Spanish equivalent (Test de Vocabulario en Imagenes Peabody) were used as pretests and as covariates in the final analyses. As expected, first graders in the TBE classes scored significantly higher on the Spanish Woodcock reading posttest (+0.60) and significantly lower in English (-0.41) than their SEI counterparts. However, the differences between the two conditions started to narrow by second and third grades. By fourth grade, no significant differences were found between conditions on all three English reading measures. Similarly, there were no significant differences in Spanish posttests. The English differences had a mean effect size of -0.26. The findings of this study suggest that, as the authors put it, “What matters most in the education of ELLs is the quality of instruction, not the language of instruction.”

Effective Reading Programs for ELLs

In the next section, we will systematically review research on effective reading program for ELLs other than use of native language.

Inclusion Criteria for Review on Effective Reading Programs for ELLs

To be included in the effective reading programs section of this review, studies had to meet the following inclusion criteria.

1. The studies involved K-12 students identified as ELLs or language minority students in English speaking countries.

2. Either random assignment to conditions was used, or pretesting or other matching criteria established the degree of comparability of the treatment and control groups. Analyses adjusted for pretest differences or data permitting such adjustments were presented. Studies without control groups, such as pre-post comparisons or comparisons to “expected” scores or gains, were excluded. Studies with pretest differences exceeding half of a standard deviation were excluded.
3. The subjects were Spanish-dominant ELLs in elementary schools in the U.S. Studies that mixed ELLs and English monolingual students in a way that did not allow for separate analyses were excluded (e.g., Hurley, Chamberlain, Slavin, & Madden, 2001; Skoczymas, 1972). Studies of children learning a foreign language were not included.
4. The language of instruction was English in both experimental and control groups.
5. The dependent measures included quantitative measures of English reading performance, such as standardized reading measures. In all cases, measures included assessment of comprehension, not just phonics or decoding. Measures of content taught in the treatment group but not the control group, such as a specific set of target words taught in a vocabulary intervention, were excluded. The focus on quantitative measures was intended to allow for comparable, objective conclusions about program effectiveness across studies.
6. A minimum treatment duration of 12 weeks was required.

Findings

Whole-School and Whole-Class Interventions

Success for All (SFA)

Among the beginning reading studies that met the inclusion criteria, three evaluated the Success for All program (Slavin & Madden, 2001; Slavin, Madden, Chambers, & Haxby, 2009). 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.

English-Language Development (ELD) Adaptation of Success for All. Ross and his colleagues (1998) conducted a one-year matched control study on the ELD adaptation of Success for All (SFA) in six schools in an Arizona school district. Participants were 540 first grade Spanish-dominant students in two SFA schools using an ELD adaptation of SFA and 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. After adjusting for initial differences, Spanish-dominant Success for All students scored significantly higher than control students on all measures, with a median effect size of +0.46.

Success for All with Embedded Video. The second study was by Chambers, Slavin, Madden, Cheung, & Gifford (2004), who investigated the effectiveness of an adaptation of Success for All that incorporated embedded video. Four types of video materials 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. Spanish-dominant students were expected to benefit in particular from the embedded video treatment because the videos included vocabulary presentations and clear, visual reinforcements of reading skills. A total of 455 K-1 Hispanic students (311 treatment and 144 control) in eight schools in New York City, Washington DC, rural Arizona, and southern California participated in this one-year long matched control study. The two groups were well-matched on their pretest scores. Analyses of covariance, using pretests as covariates, found that 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), with an overall median effect size of +0.36.

Bilingual Transition with Success for All. A one-year matched control experiment was carried out by Calderón, August, Slavin, Durán, Madden, & Cheung (2004), evaluating 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 was a modified version of Bilingual Cooperative Integrated Reading and Composition (BCIRC), which consisted of 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 et al. (2004). Participants were 238 Spanish dominant students in eight schools in El Paso, Texas. The study compared students who

received the full program to matched students in similar control schools. After controlling for Spanish and English Woodcock Scales, treatment students scored higher than control students on Woodcock Word Attack (ES=+0.21), Passage Comprehension (ES=+0.16), and Picture Vocabulary (ES=+0.11), with a median effect size of +0.16.

The effects of Success for All on the achievement of Spanish-dominant ELLs were generally positive. Across all three studies, the overall weighted effect size was +0.36.

Embedded Multimedia in Success for All. A study of the embedded multimedia component of SFA was conducted by Chambers, Cheung, Madden, Slavin, & Gifford (2004). It compared Success for All schools using the embedded video materials described above to schools also implementing Success for All but without the embedded videos. Since all ten participating schools used SFA, this was not a study of Success for All, but of the added embedded video treatment. A total of 172 first-grade Hispanic students in inner-city Hartford, CT were randomly assigned to SFA plus embedded video or SFA-only (control) conditions for a one-year experiment. Results for Spanish-dominant children, who were 66% of the sample, found positive effects controlling for the PPVT and the Woodcock Word Identification scale on Woodcock Word Identification (E=+0.23), Word Attack (ES=+0.36), and Passage Comprehension (ES=+0.16), and DIBELS Fluency (ES=+0.07), with an overall median effect size of +0.20.

Literacy Intervention with Cooperative Learning

Bilingual Cooperative Integrated Reading and Composition (BCIRC). A 5-year experiment by Calderón, Hertz-Lazarowitz, & 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, 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.

Participants were 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 (3 schools) or control (4 schools) conditions. The experimental and control groups were well matched on pretest and demographics. Two cohorts were assessed, one of which was involved for just one year (second grade) and the other for two 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, with a median effect size of +0.54.

Peer-Assisted Learning Strategies (PALS). A small matched control study was carried out by Saenz (2002) to evaluate the effectiveness of Peer-Assisted Learning Strategies (PALS) for limited English proficient students with learning disabilities and their ELL peers. A total of 132 students and 12 teachers from 12 classrooms participated in this study. Students and teachers were well matched on demographic characteristics and achievement data. The duration of the study was 15 weeks. Key components of PALS included partner reading with story retell, paragraph shrinking, prediction relay, and terms and points. Teachers in the treatment condition received a full-day workshop on PALS. Teachers in the control condition were asked to conduct their reading instruction in their normal fashion. At the conclusion of the study, significant differences were detected between the treatment and control conditions on all three measures: Words correct (ES=+0.17); questions correct (ES=+0.76), and maze choice (ES=+0.16), with an overall median effect size of +0.17.

Direct Instruction (DI)

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 English language learners 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 U.S., 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 followup of the Follow Through study when the children who had experienced the treatments in grades K-3 were in grades 5-6. Participants were 225 Hispanic English language learners. The Uvalde DI students were well matched on demographic factors with their control group. After 2 years, the treatment group scored significantly higher than the controls on both WRAT and MAT. Effect sizes averaged +0.47 for two scales of the individually administered WRAT and +0.16 across three Metropolitan Achievement Test (MAT) subscales, for a median effect size across five tests in two grades of +0.21.

Vocabulary Intervention

Instruction with Key Vocabulary. Carlo, August, McLaughlin, Snow, Dressler, Lippman, Lively, & White (2004) conducted a two-year evaluation of a vocabulary teaching intervention with 142 Spanish-dominant ELL fifth graders (94 treatment and 48 control students) 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 five-week unit and two six-week units in the first year, and three five-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 English pretests such as Peabody Picture Vocabulary Test Revised and reading comprehension. At the end of the first year, ELLs showed greater gains from pretest than controls, but surprisingly, gains were lower after two years of intervention. The median effect size across five English measures in Year 2 was +0.21.

Improving Comprehension Online (ICON). This quasi-experimental study conducted by Proctor and his colleagues (Proctor et al., 2011) evaluated the effectiveness of an internet-delivered vocabulary and comprehension intervention that targeted both English-speaking and Spanish-dominant students. The

ICON intervention was integrated into the existing curriculum and consisted of two 50-minute sessions per week for 16 weeks in a school computer lab. A total of 12 classrooms and teachers were assigned to either ICON or traditional literacy curriculum. A total of 240 fifth grade students from four schools in three primarily Hispanic districts in a northeast metropolitan area participated in the study. One hundred and eighteen of them were bilingual students (59 treatment and 59 control). The treatment students worked in a strategic digital reading (SDR) intervention that was designed to improve both vocabulary and reading comprehension in a whole class setting. The main features of SDR included: Spanish translations of all texts, human read-alouds of each text in English and Spanish; a revisable electronic worklog that collected student response, a multimedia glossary, etc. While students worked on ICON, teachers monitored and reviewed student's work in the electronic worklogs. No statistically significant differences were found for Spanish-dominant or English-dominant students. For the Spanish-dominant students, effect sizes on the Gates-MacGinitie reading achievement tests of vocabulary was +0.02 and for comprehension it was also +0.02, after adjusting for initial pretest difference.

Academic Language Instruction for All Students (ALIAS). Another study of an academic vocabulary program was carried out by Lesaux and her colleagues (Lesaux, Kieffer, Faller, & Kelley, 2010). Teachers in 21 classes were randomly assigned to treatment or control class within each school. Participants were 476 sixth grade students in seven middle schools in urban districts in California. Over 70% of them were language minority students and approximately 60% of them listed Spanish as their home language. The intervention was a text-based academic language program designed to “build knowledge of the words incrementally over time by providing multiple exposures to the words in different forms and in different meaningful context.” The program was 18 weeks long and had 8 two-week units, including an 8-day lesson cycle and 2 one-week review units. The two groups were well matched on their pretest scores. After adjusting for initial pretest difference, the treatment language minority students scored significantly higher than the control students on all four researcher-developed measures: Target word mastery (ES=+0.39, $p<0.001$), morphological decomposition (ES=+0.22, $p<0.001$), word-meanings-in-context (ES=+0.20, $p<0.05$), target word association (ES=+0.15, $p<0.08$). On the two standardized measures, the treatment group only scored slightly higher than the controls: Gates-MacGinitie reading (ES=+0.15, $p<0.06$), and SAT-10 reading vocabulary (ES=+0.01, ns). We excluded the four researcher-developed measures due to its inherence to the treatment. The effects were similar between language minority students and English-only students. The median effect size for language minority students across the two standardized measures was +0.08.

The weighted effect size for all three vocabulary interventions was +0.10.

Enhanced Language Arts Transition Program

The enhanced language arts transition program developed by Saunders and Goldenberg (1999) was designed to help English language learners transition from Spanish to English. The intervention is an enhanced version of a language arts program for grades 2-5 with 12 key instructional components in three key areas: studying literature (literature units, literature logs, instructional conversations, culminating writing projects); skill building: (comprehension strategies, assigned independent reading, dictation, written conventions lessons, ELD through literature); and other supporting components (pleasure reading, teacher read-alouds, and interactive journals). Professional development was provided by the research team twice a month for two hours after school and 3-4 times a year for a full day.

Saunders (1999) carried out a study to compare children in the three-year transition program to those in a three- to six-month transition, the usual treatment for ELLs in the district studied. Participants were 84 Spanish-dominant students from 10 schools in the same school district. The treatment and control groups were well-matched on SES, and their first grade Spanish reading and language scores. 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=+0.40). In 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=+0.58 for English, ES=+0.81 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.

English Language and Literacy Acquisition (ELLA)

A project called English Language and Literacy Acquisition (ELLA) provided students with an intervention comprised of three tiers. Tier I was the regular language arts, mathematics, science, and social studies instruction in Spanish in kindergarten and first grade. Tier II was the English intervention, including three integrated strands: 1) daily tutorial from the Santillana Intensive English program (Ventriglia & Gonzalez, 2000); 2) a storytelling and retelling activity (Irby, Lara-Alecio, Quiros, Mathes, & Rodriguez, 2004); and 3) teacher-conducted academic oral language in kindergarten and academic oral language in science in first grade. Tier III was intensive English tutorials delivered in small groups by

highly qualified paraprofessionals for low-performing students. Teachers were provided with regular professional development workshops by the research team.

Three closely-linked ELLA studies were included in this review. Irby and her colleagues (2010) conducted a 4-year longitudinal study to examine the impact of ELLA. It used a 2x2 design (treatment x language of instruction), in which classes within each language of instruction were randomly assigned to either an enhanced intervention or typical method. Participants were 381 Spanish-dominant ELLs from 22 low SES elementary schools in an urban school district in Southeast Texas. In order to be qualified to participate in the study, schools had to house programs of either SEI or TBE. Schools were randomly assigned to either treatment (enhanced) or control (typical) condition. For the SEI comparison, the results were mixed. The treatment group scored higher than the control group on two of three English measures: TAKS Reading Test (ES=+0.14) and Listening Comprehension (ES=+0.13). But the control group scored significantly higher than the treatment group on Passage Comprehension with an effect size of -0.70. The median effect size across three measures was +0.14. For the TBE comparison, the treatment group scored higher than the controls on three Spanish outcome measures with a median effect size of +0.18. Since the control group for the TBE comparison was not tested in English, no comparison was made on their English outcome measures.

A report by Tong et al. (2011) was part of the larger experimental longitudinal study mentioned above (see Irby, et al., 2010). The main focus of this study was to investigate ELLs' dual language performance in oral language and reading skills from kindergarten to the end of first grade across treatment and gender. A total of 140 students (70 in each condition) were randomly selected from the larger longitudinal study. All of these students were placed in TBE classes in 10 schools and 12 classrooms. Students were tested three times for oral language (fall kindergarten, spring kindergarten, and spring of grade 1) and two times for literacy skills (fall and spring of grade 1). Results showed that students who received the enhanced TBE program outperformed their counterparts in the control condition on only two of the six English outcome measures: IDEA Oral Language Proficiency Test (ES=+0.48) and Woodcock Passage Comprehension (ES=+0.15). The median effect size across all six English measures was +0.01. However, the treatment group scored significantly higher than the controls on 5 out of the 6 Spanish outcome measures, with a median effect size of +0.28.

The third ELLA study was a 3-year (K-2) longitudinal randomized study (Tong, Irby, Lara-Alecio, & Mathes, 2008) derived from the larger 4-year longitudinal study mentioned above. The main objective of

this study was to look at the effectiveness of ELLA in the TBE classrooms. Nineteen schools were randomly assigned to either treatment (N=10) or control schools (N=9). Treatment students received an enhanced developmental bilingual education (DBE) program, which used a 70% Spanish and 30% English late-exit model, whereas the control group used a more typical 80% Spanish and 20% English bilingual model. The initial sample size was 502 and the final sample was 262, with an attrition rate of 48%. No differential attrition rate was found between the two groups. The findings indicated that treatment students outperformed their control counterparts in the areas of oral language, pre-literacy skills, and reading fluency and comprehension on English measures with effect sizes ranging from +0.13 to +0.70. The median effect size across all 14 English measures was +0.23. The treatment students also scored significantly higher than the control students on 7 out of the 14 Spanish outcome measures with a median effect size of +0.12. Overall the ELLA produced a weighted effect size of +0.14.

Language and Literacy Curriculum

Francis and his colleagues (Francis et al., 2009) conducted a 4-year longitudinal study (K-3) to examine the effects of the Language and Literacy Curriculum. The intervention focused on developing students' literacy skills as well as their oral language proficiency skills, including listening comprehension and vocabulary and providing more practice reading connected text. The materials developed for the structured immersion program were in English whereas those developed for bilingual were in Spanish. Cooperative learning strategies were used and professional development was provided to all treatment teachers throughout the school year on a monthly basis. The study employed a 2x2 factorial design (treatment x language of instruction). Teachers within each language of instruction were randomly assigned to either the treatment or control condition. A total of 1,271 kindergarteners and 55 teachers from 13 public schools in Brownsville, Texas participated in the year one study. At third grade, the sample size was reduced to 744 students and 11 schools due to attrition. Students were assessed twice a year (fall and spring) from kindergarten to third grade. Out of the 13 original participating schools, five of them used only transitional bilingual education, one school used only the structured immersion program, and the other seven had both SEI and TBE programs. Results were mixed. For the SEI group, after adjusting for pretest differences, the treatment group scored lower than control group on Basic Reading (ES=-0.18), Broad Reading (ES=-0.26), and Oral Language Composition (ES=-0.51), with a median effect size of -0.26 across all three English measures. The treatment group scored significantly higher than the control group on all three Spanish measures with a median effect size of +0.58. For the bilingual group, the treatment students only scored slightly higher than the controls on two of three

English measures with a median effect size of +0.07 but significantly higher on all three Spanish measures with a median effect size of +0.81. The weighted effect size for the English measures for all ELLs was -0.10.

Small-Group and One-to-One Supplemental Interventions

Small Group Emergent Literacy Intervention

Small Group Tutorials with Direct Instruction. A randomized experiment was carried out by Gunn, Biglan, Smoklowski, & Ary (2000) to evaluate the effects of a small group supplemental tutorial program that used two forms of DI, *Reading Mastery* and *Corrective Reading*, for Hispanic and non-Hispanic children who were struggling in reading. Participants were at-risk early elementary school children (K-3) and were selected from nine rural Oregon elementary schools in three school districts. After screening for eligibility, 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 two 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 well matched on their pretest scores. After the first year, the treatment students outperformed control students on all three measures, Letter-Word ID (ES=+0.22), Word Attack (ES=+0.70), and Fluency (ES=+0.16), with an overall mean effect size of +0.36. 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. The median effect size across these three measures was +0.46. 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 two years, with a median effect size of +0.46 on the English measures.

An Explicit, Systematic Supplemental Reading Intervention. Vaughn, Mathes, Linan-Thompson, & Francis (2005) evaluated a supplemental program in which trained bilingual teachers provide systematic and explicit instruction in phonemic awareness and phonics applied to word and text reading in English. Students meet in small groups (3-5) for 50 minutes daily, five days a week. Extensive professional development is provided to teachers before and during the intervention.

Vaughn, Mathes, and colleagues (2006) carried out a small-scale randomized study examining the effectiveness of this intervention for at-risk first grade ELLs. To be eligible for the study, students needed to score below the 25th percentile on the Woodcock Letter-Word Identification pretests in both English and Spanish at the beginning of first grade. Out of the 216 students in four low SES schools from two districts in Texas, 56 students were eligible for the study. The study duration was from October to May. The intervention was provided in addition to the core reading lessons. Students in the control classroom also received one or more types of supplemental reading intervention in addition to their core reading instruction. The treatment and control students were well-matched on both English and Spanish pretest measures. At the end of the 7-month study, treatment students scored significantly higher than control students on 7 of the 14 English outcome measures. Strongest effects were in the areas of phonemic awareness (ES=+1.24), Word Attack (ES=+1.09), Passage Comprehension (ES=+1.08), and phonological processing (ES=+1.01). The Spanish measures were also in favor of the treatment group. The median effect size for all 14 English measures was +0.68.

Vaughn, Cirino, and colleagues (2006) replicated a pair of similar randomized experiments with two separate samples of at risk first grade ELLs to investigate the effectiveness of both the English and Spanish intervention as reported in the two previous studies (Vaughn, Linan-Thompson, et al., 2006; Vaughn, Mathes et al., 2006). Due to the focus of this review, only the study that used English outcome measures will be reported here. Participants were 90 students from four schools and 20 classrooms from three sites in Texas. Students in the study were first screened and then randomly assigned to either the treatment or the control condition (using the methods described in Vaughn, Mathes, et al., 2006). Treatment students were provided about 115 sessions of supplemental reading for 50 minutes every day in a small group setting (3-5 students). At the end of the study, treatment students outperformed their counterparts in the control condition in phonological awareness, word attack, word reading, and spelling, with effect sizes ranging from +0.40 to +0.75. The median effect size across all 16 English measures was +0.27.

Literacy Express Preschool Curriculum. An evaluation was conducted by Farver and her colleagues (Farver, Lonigan, & Eppe, 2009) to examine the effects of the Literacy Express Preschool Curriculum that provides intensive small group instruction to ELLs. A total of 94 Spanish-dominant ELL preschoolers from a Head Start preschool program in Los Angeles participated in this 6-month long randomized study. Students were randomly assigned to one of the three conditions: High/Scope

Curriculum (Control=32), the Literacy Express Preschool Curriculum in English-only (Treatment English=31), and the Literacy Express Preschool Curriculum initially in Spanish then transitioning to English (Treatment Transition=31).

The Literacy Express Preschool Curriculum is a comprehensive preschool program designed to improve young children's oral language, emergent literacy skills, and socio-emotional development. The curriculum is structured around 10 thematic units and provides intensive small group (4-5 students) instruction that focuses on dialogic reading activities, phonological awareness activities, and print knowledge activities. During the duration of the study, both treatment groups received a 20-minute small group intervention four times a week for six months. The first treatment group received the intervention from the Literacy Express Preschool Curriculum in English. The transition group received the same intervention in Spanish for two months and was then transitioned to English instruction. The transition period took place over 3-4 weeks. The treatment groups and control group were well matched on pretest measures. Analyses of covariance indicated that both the English treatment and transition treatment group outperformed the control group on all five English measures with median effect sizes of +0.41 and +0.71, respectively, with a combined effect size of +0.56. The transition treatment group also scored significantly higher than the control group on all five Spanish language outcomes. No significant difference was found on Spanish outcome measures between the English treatment and control group.

Across the four small-group supplemental studies, the weighted effect size was +0.47.

One-on-One Tutoring: *Read Well* and *Read Naturally*

Read Naturally is a supplementary reading program aimed at improving reading fluency using a combination of books, audiotapes, and computer software for elementary and middle school students. Key strategies used include repeated reading of text for developing oral reading fluency, teacher modeling of story reading, and systematic monitoring of student progress by teachers and the students themselves (Ihnot, 1992).

Read Well combines systematic, explicit phonics instruction with practice in decodable text and contextualized vocabulary and comprehension instruction. Activities in the *Read Well* program include tutor-directed decoding practice, practice reading decodable text with pre-reading and during-reading discussion and questioning designed to build vocabulary and comprehension, and completion of simple

comprehension worksheets (Sprick, Howard, & Fidanque, 1998).

Read Well and Read Naturally. Two types of one-to-one tutoring for English language learners were studied in an experiment by Denton, Anthony, Parker, & 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 *Read Well* 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. 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 ($p < .016$) 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 ($ES = +0.08$).

Read Naturally. Kemp (2006) carried out a randomized experiment to examine how *Read Naturally* influenced reading and reading-related skills. Forty-two third grade ELLs from three schools were randomly assigned to *Read Naturally* or the control condition. All students participated in 20 minutes of independent reading four days a week. The treatment group received *Read Naturally* whereas the control group received scaffolded sustained silent reading for the same amount of time. After the 4-month intervention, the two groups did not differ significantly on all five outcome measures, with an overall median effect size of +0.10.

Phonics-based supplemental reading intervention. Vadasy and Sanders (Vadasy & Sanders, 2011) evaluated a supplemental phonics-based tutoring intervention. Participants included both language minority students and non-language minority students. Due to the focus of this review, only results from the language minority students are reported here. Ninety-eight first graders who performed in the lower half of their class were randomly assigned to either treatment ($N=48$) or control condition ($N=50$) within each classroom. The treatment students received 30 minutes of individual tutoring in English for 4 days a week for 6 months. Scripted lessons were provided to each tutor and each lesson focused several key components including letter-sound correspondences, phoneme decoding, irregular words, spelling, and

oral reading practice. Tutors also received ongoing coaching and modeling of appropriate scaffolding from research staff. Treatment and control students were fairly well matched on key demographic characteristics and pretest scores. The majority of the participants were Spanish-dominant students. After the intervention, the treatment students scored higher than the control students on all outcome measures, with effect sizes ranging from +0.03 to +0.40. The median effect size across all six measures was +0.22. Overall, the three one-to-one tutoring programs produced a combined weighted effect size of +0.21.

Discussion

The purpose of this review was to synthesize research on outcomes of all types of programs likely to improve English reading outcomes for Spanish-dominant English language learners in elementary school. In the past, the focus in this area has been on language of instruction, with reviewers debating the merits of bilingual versus English-only approaches. In the present review, however, we have treated various forms of bilingual education as “interventions” for ELLs, and have considered their outcomes on the same basis as we consider other types of interventions intended to improve English reading outcomes for ELLs. This is the first systematic review to consider all alternatives together in this way, focusing on studies that meet high, consistent standards of methodological rigor.

The findings of the review support a conclusion increasingly being made by researchers and policy makers concerned with optimal outcomes for ELLs and other language minority students: Quality of instruction is more important than language of instruction. Combining outcomes from 14 qualifying studies of bilingual education going back to the 1970s, average effect sizes weighted by sample size favor bilingual education, with an effect size of +0.19. However, the largest and longest-term evaluations, including the only multiyear randomized evaluation of transitional bilingual education, did not find any differences in outcomes by the end of elementary school for children who were either taught in Spanish and transitioned to English or taught only in English.

More importantly, there are several types of interventions that have been found to be more effective in improving reading outcomes for Spanish-dominant ELLs. Some of these are widely used and capable of being used broadly. For example, Success for All, a whole-school reform approach with specific adaptations for English language development, was found to have positive effects with a weighted mean effect size of +0.36 across 3 studies. Two forms of cooperative learning had positive effects on ELLs;

one, Bilingual Cooperative Integrated Reading and Composition (BCIRC), had an effect size of +0.54, and the other, Peer Assisted Learning Strategies, had an effect size of +0.17 for ELLs. Positive effects were also reported for Direct Instruction and ELLA. Success for All, cooperative learning, DI, and ELLA are all cost-effective whole-class or whole-school interventions that exist at some scale in the U.S. and could be introduced in any school serving many ELLs. Other whole-class interventions evaluated by Carlo et al. (2004), Leseaux et al. (2010), Francis et al. (2011), and Saunders (1999), are not in wide use, but they still illustrate the potential to make significant improvements in the English reading of ELLs.

Another category of promising and scalable interventions includes small-group and one-to-one tutoring for English language learners who are struggling in reading. Programs of this kind focusing on phonics and language development have shown great promise, especially structured small-group programs evaluated by Vaughn et al. (2006 a, b), Gunn et al. (2000), and Farver et al. (2009).

Looking across the most promising interventions, there are several themes that appear. First, all of the proven programs provide extensive professional development and coaching to help teachers effectively implement promising models. Effective programs provide explicit manuals, videos, and simulations to start teachers off in the right direction, and then have experienced coaches visit teachers using new strategies to offer feedback and support. Second, almost all of the effective strategies make extensive use of cooperative learning, which gives English language learners extensive, daily opportunities to use their developing language skills in meaningful contexts. In particular, it seems likely that cooperative learning is essential for building confidence in the use of school-specific English, of the kind students are unlikely to hear on the playground or in their communities.

Language of instruction remains an important question, if for no other reason than that building on students' home language gives them skills in that language that are important in their lives. However, when English reading is the goal, different approaches may work equally well, bilingual as well as structured English immersion, but we now have many other approaches with evidence of effectiveness from rigorous evaluations. Educators and policy makers should consider all possibilities to enhance outcomes for their Spanish-dominant ELL children.

*References marked with an asterisk indicate studies included in the review

References

- Adams, G. L., & Engelmann, S. (1996). *Research on Direct Instruction: 25 years beyond DISTAR*. Seattle, WA: Education Achievement Systems.
- *Alvarez, J. (1975). *Comparison of academic aspirations and achievement in bilingual versus monolingual classrooms*. Doctoral dissertation, University of Texas, Austin, TX.
- August, D., & Hakuta, K. (1997). *Improving schooling for language-minority children: A research agenda*. Washington, DC: National Research Council.
- *Barnett, W. S., Yarosz, D. J., Thomas, J., Jung, K., & Blanco, D. (2007). Two-way and monolingual English immersion in preschool education: An experimental comparison. *Early Childhood Research Quarterly*, 22(3), 277-293.
- *Becker, W. C., & Gersten, R. (1982). A follow-up of the follow through: The latter effects of the direct instruction model on children in fifth and sixth Grades. *American Educational Research Journal*, 19(1), 75-92.
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2005). *Comprehensive meta-analysis (Version 2)*. Englewood, NJ: Biostat.
- *Calderon, M., August, D., Slavin, R. E., Duran, D., Madden, N. A., & Cheung, A. (2004). The evaluation of a bilingual transition program for Success for All. Baltimore, MD, Center for Research on the Education of Students Placed at Risk.
- Calderon, 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.
- *Campeau, P. L., Roberts, A., Oscar, H., Bowers, J. E., Austin, M., & Roberts, S. J. (1975). *The identification and description of exemplary bilingual education programs*. Palo Alto, CA: American Institutes for Research.
- *Carlo, M. S., August, D., McLaughlin, B., Snow, C. E., Dressler, C., Lippman, D. N., . . . White, C. E. (2004). Closing the gap: Addressing the vocabulary needs of English-Language learners in bilingual and mainstream classrooms. *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, Maryland: Johns Hopkins University, Center for research on the Education of Students Placed at High 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, MD: Johns Hopkins University, Center for Research on the Education of Students Placed at Risk.

- Child Trends Data Bank. (2011). High school dropout rates: Indicators on children and youth. Retrieved November 8, 2011 http://www.childtrendsdatabank.org/sites/default/files/01_Dropout_Rates.pdf
- Christain, D., & Genesee, F. (2001). *Bilingual Education*. Alexandria, VA: TESOL.
- *Cohen, A. D. (1975). *A sociolinguistic approach to bilingual education*. Rowley, MA: Newbury House Press.
- Cuirel, H., Stenning, W., & Cooper-Stenning, P. (1980). Achieved ready level, self-esteem, and grades as related to length of exposure to bilingual education. *Hispanic Journal of Behavioral Sciences*, 2, 389-400.
- Danoff, M. N., Coles, G. J., McLaughlin, D. H., & Reynolds, D. J. (1978). Evaluation of the impact of ESEA Title VII Spanish/ English Bilingual Education Programs: Vol.I. 3. Year Two impact designs. Palo Alto, CA: American Institutes for Research.
- *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.
- *Farver, J. A. M., Lonigan, C. J., & Eppe, S. (2009). Effective early literacy skill development for young Spanish-Speaking English language learners: An experimental study of two methods. *Child Development*, 80(3), 703-719.
- Francis, D., & Vaughn, S. (2009). Effective practices for English language learners in the middle grades: Introduction to the special of *Journal of Research on Educational Effectiveness*. *Journal of Research on Educational Effectiveness*, 2(4), 289-296.
- *Francis, D., York, M., August, D., & Vaughn, S. (2009). Report on the effectiveness of the Language and Literacy Curriculum. Washington, D.C.: U.S. Department of Education.
- Genesee, F. (Ed.). (1999). *Program alternatives for linguistically diverse students*. Santa Cruz, CA: Center for Research on Education, Diversity, and Excellence.
- Genesee, F., Paradis, J., & Crago, M. (2004). *Dual language development and disorders: A handbook on bilingualism and second language learning*. Baltimore, MD: Brookes.
- 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 Rossel & 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.
- Guo, Q. (in press). Estimating the Impact of the Massachusetts English Immersion Law on Limited English Proficient Students' Reading Achievement. *Educational Policy*.

- 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.
- *Huzar, H. (1973). *The effects of an English-Spanish primary grade reading program on second and third grade students*. Master's thesis, Rutgers University, Camden, NJ.
- lnhot, C. (1992). *Read Naturally*. St. Paul, MN: Read Naturally.
- *Irby, B. J., Tong, F., Lara-Alecio, R., Mathes, P. G., Acosta, S., & Guerrero, C. (2010). Quality of instruction, language of instruction, and Spanish-speaking English language learners' performance on a state reading achievement test. *TABE*, 12(1), 1-42.
- Irby, B. J., Lara-Alecio, R., Quiros, A. M., Mathes, P. G., & Rodriguez, L. (2004). English language and literacy acquisition evaluation research program (Project ELLA): Second annual evaluation report. Washington, DC: U.S. Department of Education, Institute of Educational Sciences.
- *Kemp, S. C. (2006). *Teaching to Read Naturally: Examination of a fluency training program for third grade students*. Doctoral dissertation, University of California, Irvine and University of California, Los Angeles.
- *Lesaux, N. K., Kieffer, M. J., Faller, S. E., & Kelley, J. G. (2010). The effectiveness and ease of implementation of an academic vocabulary intervention for linguistically diverse students in urban middle schools. *Reading Research Quarterly*, 45(2), 196-228.
- Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.
- *Maldonado, J. R. (1977). *The effect of the ESEA Title VII program on the cognitive development of Mexican American students*. Doctoral dissertation, University of Houston, Houston, TX.
- Meyer, M. M., & Fienberg, S. E. (1992). Assessing evaluation studies: The case of bilingual education strategies. Washington, DC: National Academy of Sciences.
- Morgan, J. C. (1971). *The effects of bilingual instruction of the English language arts achievement of first grade children*. Doctoral dissertation, Northwestern State University of Louisiana, Natchitoches, LA.
- National Assessment of Educational Progress. (2011). The Nation's report card. Retrieved November 1, 2011 http://nationsreportcard.gov/reading_2011/nat_g4.asp?tab_id=tab2&subtab_id=Tab_7#chart
- National Clearinghouse for English Language Acquisition & Language Acquisition, T. (2011). The growing numbers of English learner students. Retrieved November 4, 2011 http://www.ncela.gwu.edu/files/uploads/9/growingLEP_0809.pdf
- Parrish, T. B., Merickel, A., Perez, M., Linqanti, R., Socias, M., Spain, A., . . . Delancey, D. (2006). Effects of the implementation of Proposition 227 on the education of English learners, K-12: Findings from a five year evaluation. Palo Alto, CA: American Institute for Research.
- *Plante, A. J. (1976). A study of effectiveness of the connecticut "pairing" model of bilingual/bicultural education. Hamden, CT: Connecticut Staff Development Cooperative.

- *Proctor, C. P., Dalton, B., Uccelli, P., Biancarosa, G., Mo, E., Snow, C., & Neugebauer, S. (2011). Improving comprehension online: Effects of deep vocabulary instruction with bilingual and monolingual fifth graders. *Reading and Writing, 24*(5), 517-544.
- *Ramirez, J., Pasta, D. J., Yuen, S., Billings, D. K., & Ramey, D. R. (1991). Final report: Longitudinal study of structural immersions strategy, early-exit, and late-exit transitional bilingual education programs for language-minority children. San Mateo, CA: Aguirre International.
- *Ross, S. M., Smith, L. J., & Nunnery, J. A. (1998). *The relationship of program implementation quality and student achievement*. Paper presented at the American Educational Research Association San Diego, CA.
- Rossell, C., & Baker, K. (1996). The educational effectiveness of bilingual education. *Research in the Teaching of English, 30*(1), 7-69.
- *Saenz, L. M. (2002). *Peer-assisted Learning Strategies for limited English proficient students with learning disabilities*. Doctor of Philosophy, Graduate School of Vanderbilt University, Nashville.
- *Saldate, M., Mishra, S. P., & Medina, M. (1985). Bilingual instruction and academic achievement: A longitudinal study. *Journal of Instructional Psychology, 12*(1), 24-30.
- *Saunders, W. M. (1999). Improving literacy achievement for English learners in transitional bilingual programs. *Educational Research and Evaluation, 5*(4), 345-381.
- Saunders, W. M., & Goldenberg, C. (1999). The effects of a comprehensive language arts transition program on the literacy development of English language learners. Long beach, CA: Center for Research on Education, Diversity, and Excellence, University of California.
- Sedlmeier, P., & Gigerenzer, G. (1989). Do studies of statistical power have an effect on the power of studies? *Psychological Bulletin, 105*, 309-316.
- Skoczylas, R. V. (1972). *An evaluation of some cognitive and affective aspects of a Spanish bilingual education program*. doctoral dissertation, University of New Mexico, Albuquerque, NM.
- Slavin, R. E. (1986). Best-evidence synthesis: An alternative to meta-analytic and traditional reviews. *Educational Researcher, 15*(9), 5-11.
- Slavin, R. E. (2008). Perspective on evidence based research in education: What works? Issues in synthesizing education program evaluations. *Educational Researcher, 37*(1), 5-14.
- Slavin, R. E., & Cheung, A. (2005). A synthesis of research on language of reading instruction for English language learners. *Review of Educational Research, 75*(2), 247-284.
- Slavin, R. E., & Madden, N. A. (2001). *One million children: Success for All*. Thousand Oaks, CA: Corwin.
- *Slavin, R. E., Madden, N. A., Calderon, M., Chamberlain, A., & Hennessy, M. (2011). Reading and language outcomes of a multiyear randomized evaluation of transitional bilingual education. *Education Evaluation and Policy Analysis, 33*(1), 47-58.

- Slavin, R. E., Madden, N. A., Chambers, B., & Haxby, B. (2009). *2 million children: Success for All* (2nd ed.). Thousand Oaks, CA: Corwin Press.
- Sprick, M. M., Howard, L. M., & Fidanque, A. (1998). *Read well: Critical foundations in primary reading*. Longmont, CO: Sporis West.
- Stebbins, L. B., St. Pierre, R. G., Proper, E. C., Anderson, R. B., & Cerva, T. (1977). *Education as experimentation: A planned variation model. IV-A*. 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.
- Thomas, W. P., & Collier, V. P. (1997). *School effectiveness for language minority students*. Washington, D.C.: National Clearinghouse for Bilingual Education.
- Thomas, W. P., & Collier, V. P. (2002). *A national study of school effectiveness for language minority students' long-term academic achievement*. Santa Cruz, CA: Center for Research on Education, Diversity, and Excellence, University of California at Santa Cruz.
- *Tong, F., Irby, B. J., Lara-Alecio, R., & Mathes, P. G. (2008). English and Spanish acquisition by Hispanic second graders in developmental bilingual programs. *Hispanic Journal of Behavioral Sciences*, 45(4), 1011-1044.
- *Tong, F., Lara-Alecio, R., Irby, B. J., & Mathes, P. G. (2011). The Effects of an instructional intervention on dual language development among first-grade Hispanic English-learning boys and girls: A two-year longitudinal study. *Journal of Educational Research*, 104, 87-99.
- *Vadasy, P. F., & Sanders, E. A. (2011). Efficacy of supplemental phonics-based instruction for Low-Skilled First Graders: How Language Minority Status and Pretest Characteristics Moderate Treatment Response. *Scientific Studies of Reading*, 15(6), 471-497.
- *Vaughn, S., Cirino, P. T., Linan-Thompson, S., Mathes, P. G., Cardenas-Hagan, E., Carlson, C. D., et al. (2006a). Effectiveness of a Spanish intervention and an English intervention for English-language learners at risk for reading problems. *American Educational Research Journal*, 43(3), 449-487.
- Vaughn, S., Linan-Thompson, S., Mathes, P. G., Cirino, P. T., Carlson, C. D., Pollard-Durodola, S. D., et al. (2006). Effectiveness of Spanish intervention for first-grade English language learners at risk for reading difficulties. *Journal of Learning Disabilities*, 39(1), 56-73.
- *Vaughn, S., Mathes, P. G., Linan-Thompson, S., Cirino, P. T., Carlson, C. D., Pollard-Durodola, S. D., et al. (2006b). Effectiveness of an English intervention for first-grade English language learners at risk for reading problems. *The Elementary School Journal*, 107(2), 154-180.
- Ventriglia, L., & Gonzalez, L. (2000). *Santillana Intensive English*. Miami, FL: Santillana.
- Willig, A. (1987). Examining bilingual education research through meta-analysis and narrative review: A response to Baker. *Review of Educational Research*, 57(3), 363-376.

Wright, W. E. (2007). Heritage language problems in the era of English-only and No Child Left Behind. *Heritage Language Journal*, 5(1), 1-26.

TABLE 1
Language of Reading Instruction: Descriptive Information and Effect Sizes for Qualifying Studies

Study	Intervention description	Design	Duration	N	Grade	Sample Characteristics	Evidence of Initial Equality	Posttest	Effect Size	Final ES
Studies of Paired Bilingual Education										
Plante (1976)	Paired bilingual	Randomized	2 yrs	55	1-2, 2-3	Spanish-dominant Puerto Rican students in New Haven, CT	Well matched on Spanish oral vocabulary but C>E in English pretest	English Inter-American Series		
								2nd grade	+0.78	+0.26
								3rd grade	+0.26	
Huzar (1973)	Paired bilingual	Randomized	2 & 3 yrs	160	1-2, 1-3	Disadvantaged Puerto Rican students in Perth Amboy, NJ	Well matched on IQ, SES, and initial achievement	English Inter-American Series		
								2nd grade	+0.01	+0.16
								3rd grade	0.31	
Campeau et al. (1975)-Corpus Christi	Paired bilingual	Matched	2 yrs	171	K-1	Spanish dominant students in Corpus Christi, Texas	Matched on English and Spanish pretests	English Inter-American Series	+0.45	+0.45
Ramirez et al. (1991)	Paired bilingual	Matched	4 yrs	153	K-1	Spanish dominant LEP students	very well matched on SES and home backgrounds.	English CTBS	+0.53	+0.53
Campeau et al. (1975)-Houston	Paired bilingual	Matched	3 yrs	206	K-2	Spanish dominant students in Houston, TX	Matched on language, SES, and academic achievement	English Inter-American Series	+0.54	+0.54
Maldonado (1977)	Paired bilingual	Matched	5 yrs	126	1-5	Spanish dominant students in six elementary school in Corpus Christi, TX	Matched on SES and number of years in schools	English (SRAAS)		
								2nd	+0.10	+0.09
								3rd	+0.18	
								4th	+0.07	
	5th	+0.09								
Alvarez (1975)	Paired bilingual	Matched	2 yrs	147	2	Spanish dominant children in two schools in Austin Texas	Matched on SES and initial language proficiency	California Achievement Tests		
								English reading vocab	+0.08	-0.03
								English reading comp	+0.15	
Campeau et al. (1975)-Kingsville, TX	Paired bilingual	Matched	1 yr	89	K	Spanish dominant students in Kingsville, TX	Matched on SES and ethnic mix	English Inter-American Series	+0.42	+0.42
Campeau et al. (1975)-Santa Fe	Paired bilingual	Matched	1 yr	77	1	Hispanic students in Santa Fe, New Mexico	Similar on pretests, but PBE>E	English MAT	+0.03	+0.03

Study	Intervention description	Design	Duration	N	Grade	Sample Characteristics	Evidence of Initial Equality	Posttest	Effect Size	Final ES
Studies of Two-Way Bilingual										
Cohen (1975)	Two-way bilingual	Matched	2-3 yrs	90	K-1, 1-2, 1-3	Spanish dominant students in Redwood city, CA	Matched on SES and initial language proficiency	English Inter-American Series		
								Cohort 1	0.00	0.00
								Cohort 2	0.00	
Barnett et al. (2007)	Two-way bilingual	Matched	1 yr	131	K	Spanish dominant students in a Northeastern city	Similar on English pretest	PPVT	+0.02	-0.03
								WJ Picture Vocab	-0.16	
								Phoneme	+0.22	
								Letter ID	-0.03	
								Rhyme Recognition	-0.16	
Studies of Transitional Bilingual Education										
Slavin et al. (2011)	Bilingual-5-year transition	Randomized	5 yrs	115	K-4	Spanish dominant students in six elementary school in six states	Similar on English pretests	Woodcock		
								1st grade	-0.41	-0.26
								2nd grade	-0.19	
								3rd grade	-0.19	
								4th grade	-0.26	
Campeau et al. (1975)-Alice, TX	Bilingual-1-year transition	Matched	2 yrs	125	K-1	Spanish dominant students in Alice ISD, Texas	Similar on English pretests but E>C on Spanish pretest	English Inter-American Series	+0.49	+0.49
Studies of Bilingual Education (Unspecified)										
Saldate et al. (1985)	Unspecified	Matched	3 yrs	38	1-3	Spanish dominant students in Douglas, AZ	Well matched on pretests	English tests		
								WRAT (3rd grade)	+0.89	+0.89

TABLE 2

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

Study	Intervention Description	Design	Duration	N	Grade	Sample Characteristics	Evidence of Initial Equality	Posttest	Effect Size	Median ES
Whole-School and Whole-Class Intervention										
Success For all										
Ross et al. (1998)	SFA-English Language Development Adaption	Matched	1 yr	540 in 6 schools	1	Tucson, Arizona: 39% Hispanic, 67% free lunch	Well matched on demographics and pretests	Eng. Woodcock Word ID +0.51 Word Attack +0.83 Passage Comp +0.41 Durrell +0.32		+0.48
Chambers et al. (2004a)	SFA with embedded video	Matched	1 yr	455 in 8 schools	K-1	Hispanic students in New York City, Washington, DC, rural AZ, southern CA	Well matched on PPVT	Eng. Woodcock Word ID +0.40 Word Attack +0.36 Passage Comp +0.21		+0.38
Calderon et al. (2004b)	Success for All with enriched transition	Matched	1 yr.	239 in 8 schools	3	Spanish-dominant students in El Paso, TX	Well matched on English and Spanish Woodcock measures	English Picture Vocab +0.11 Passage Comp +0.16 Word Attack +0.21		+0.16
Embedded Multimedia										
Chambers et al. (2004c)	Embedded video (SFA with embedded video vs. SFA)	Randomized	1 yr	172 in 10 schools	1	Hispanic students in Hartford, CT	Well matched on PPVT, Word ID	Eng. Woodcock Word ID +0.23 Word Attack +0.36 Passage Comp +0.16 DIBELS Fluency +0.07		+0.20
Literacy Intervention with Cooperative Learning										
Calderon et al. (1998)	Bilingual Cooperative Integrated Reading & Composition (BCIRC)	Matched	2 yrs	222	2-3	Spanish-dominant students in El Paso, TX	Well matched on demographics and pretests	Eng. TAAS Grade 3 Reading +0.54 Eng. TAAS 2 yrs Reading +0.87 Eng. TAAS 1 yr Reading +0.33		+0.54
Saenz (2002)	Peer-Assisted Learning Strategies (PALS)	Matched	15 wks	132	3-6	Spanish-dominant students in TX	Well matched on demographics and achievement data	Word Correct +0.17 Question Correct +0.76 Maze Choice +0.16		+0.17
Direct Instruction										
Becker & Gersten (1982)	Direct Instruction (DI)	Matched	follow up study--2 yrs after the treatment	225	K-3	Hispanic ELL students in Uvalde, TX	Well matched on demographics	Eng. WRAT Reading Across 2 grades Level II +0.44 Level I +0.50 Mean +0.47 Eng. MAT Word Knowledge +0.11 Reading +0.21 Total Reading +0.16 Mean +0.16		+0.21

Study	Intervention Description	Design	Duration	N	Grade	Sample Characteristics	Evidence of Initial Equality	Posttest	Effect Size	Median ES				
Vocabulary Interventions														
Carlo et al. (2004)	Direct instruction in key vocabulary	Matched	2 yrs	~130	4 & 5	ELL students in CA, VA, and MA	Well matched on pretests	Eng Vocab Assessment		+0.21				
								PPVT	-0.08					
								Polysemy prod	+0.33					
								Morphology	+0.22					
								Semantic Association	+0.21					
Eng Reading Comp	+0.17													
Proctor et al. (2011)	Improving Comprehension Online (ICON)	Matched	4 mos	118	5	Spanish-speaking students in a northeast metropolitan area	Well-matched on pretests	Gates-MacGinitie		+0.02				
								Vocabulary	+0.02					
								Comprehension	+0.02					
Lesaux et al. (2010)	Academic Language Instruction for All Students (ALIAS)	Matched	1 yr	240	6	Spanish-dominant students from an urban district in the southwestern US	Well matched on pretests	Gates-MacGinitie		+0.08				
								Reading	+0.15					
								SAT -10						
								Reading Vocab	+0.01					
Enhanced Language Arts Program														
Saunders (1999)	Enriched transition	Matched	1 yr	84	2-5	Spanish-dominant students in Southern CA	Well matched on SES and Spanish pretests	Standardized Measures of Reading		+0.58				
								Grade 1: Spanish	-0.02					
								Grade 2: Spanish	+0.26					
								Grade 3: Spanish	+0.38					
								Grade 4: Spanish	+0.40					
Grade 5: English	+0.58													
Project English Language and Literacy Acquisition (ELLA)														
Irby et al. (2010); Tong et al. (2011); Tong et al. (2008)	English Language and Literacy Acquisition (ELLA)	Matched	4 yrs	381	K-3	Spanish-dominant ELLs from an urban district in Southeast TX	Well matched on SES and pretests	TAKS Reading	+0.14	+0.14				
								Listening Comp	+0.13					
								Passage Com	-0.70					
		Matched	2 yrs	140	K-1							Oral Language	+0.48	+0.01
												Picture Vocab	-0.18	
												Listening Comp	-0.25	
		Matched	3 yrs	262	K-2							Letter Word	+0.05	
												Word Attack	-0.02	
												Passage Comp	+0.15	
								Across all 14 English measures	+0.23	+0.23				

Study	Intervention Description	Design	Duration	N	Grade	Sample Characteristics	Evidence of Initial Equality	Posttest	Effect Size	Median ES
Language and Literacy Curriculum										
Francis et al. (2009)	Language and Literacy Curriculum	Matched	4 yrs	286	K-3	Spanish-dominant students in TX	Well-matched on pretests	SEI Group		-0.10
								Basic Reading	-0.18	
								Broad reading	-0.26	
				Oral Language				-0.51		
				Median				-0.26		
				Bilingual Group						
				Basic Reading				+0.20		
Broad reading	+0.07									
Oral Language	-0.18									
Median	+0.07									
Small Group and One-to-One Supplemental Intervention										
Small Group Emergent Literacy Intervention										
Gunn et al. (2000)	Small group tutoring using Direct Instruction	Randomized	2 yrs	122	K-4	Low-achieving Hispanic students in rural Oregon	Well-matched on English Woodcock-Johnson, oral reading fluency	Eng. Woodcock		+0.46
								Letter Word	+0.46	
								Word Attack	+0.91	
								Oral Reading	+0.43	
								Vocabulary	+0.44	
Comprehension	+0.48									
Vaughn et al. (2006a)	Enhanced bilingual program	Randomized	1 yr	48	1	Low performing ELLs in TX	Well-matched on both English and Spanish pretests	Across all 14 Eng. measures	+0.68	+0.68
Vaughn et al. (2006b)	Enhanced bilingual program (English)-small group	Randomized	1 yr	90	1	Spanish-dominant students from three sites in TX	Well matched on pretests	Across all 16 Eng. measures	+0.27	+0.27
Farver et al. (2009)	Literacy Express Preschool Curriculum	Randomized	6 mos	94	PreK	Spanish-dominant ELL preschoolers in LA	Well matched on pretests	English only		+0.56
								Across five Eng. measures	+0.41	
								Transitional		
								Across five Eng. measures	+0.71	

Study	Intervention Description	Design	Duration	N	Grade	Sample Characteristics	Evidence of Initial Equality	Posttest	Effect Size	Median ES
One-to-One Tutoring										
Denton et al. (2004)	Read Well (Tutoring using systematic phonics)	Randomized	4 mos	33	2-5	Spanish-dominant bilingual students in Texas	Well matched on WRMT pretests ; E>C, ES=+0.32 (0.3<p<0.6)	English--Read Well		+0.51
								Word ID	+0.55	
								Word Attack	+0.46	
								Passage Comp	+0.00	
								Fluency	+0.18	
Accuracy	+0.78									
Comprehension	+0.82									
Kemp (2006)	Read Naturally (Tutoring using repeated readings)	Randomized	4 mos	60	3	Spanish-dominant students in western Orange county in CA	Well matched on WRMT pretests	English--Read Naturally		+0.08
								Word ID	-0.05	
								Word Attack	-0.13	
								Passage Comp	+0.16	
								Fluency	+0.23	
Accuracy	+0.30									
Kemp (2006)	Read Naturally	Randomized	4 mos	42	3	Spanish-dominant students in western Orange county in CA	Well matched on pretests	Phonemic	+0.11	+0.10
								Reading Fluency	+0.19	
								Comprehension	+0.10	
								Vocabulary	+0.08	
								Alphabets	+0.40	
Vadasy & Sanders (2011)	Phonics-based tutoring intervention	Matched	6 mos	98	1	Language minority students from 13 urban public schools	Well matched on SES and pretests	Phono Aware	+0.03	+0.22
								Word Reading	+0.28	
								Spelling	+0.24	
								Fluency	+0.19	
								Comprehension	+0.11	