

# **Effects of a Two-Way Bilingual Program on the Literacy Development of Students in Kindergarten and First Grade**

María G. López and Abbas Tashakkori  
Florida International University

## **Abstract**

This study examines the short-term effects of a two-way bilingual education program on the literacy development of students in kindergarten and first grade. This study compared the literacy development of two groups of children who received different proportions of instruction in English and Spanish. Both the experimental and control groups consisted of students who were English language learners (ELLs), as well as students who were not ELLs (i.e., native English speakers and former ELLs who reached proficiency). Students in the experimental group participated in a two-way bilingual program in which instruction was in English approximately 70% of the time and in Spanish approximately 30% of the time. Students in the control group attended the same school but were in mainstream classes and received instruction in English about 90% of the time. Comparison of test scores (on district-developed assessments and the Scholastic Reading Inventory) of the two groups at the end of the school year indicated that in spite of the initially lower pretest scores in some subject areas, the experimental group's achievement was very close to that of the control group. It is concluded that two-way bilingual programs can assist schools in improving ELLs' academic achievement in English language arts.

## **Introduction**

Perhaps one of the most indelible images of the United States is that of a nation of immigrants. Individuals have come to this country for various reasons and have brought with them different traditions, cultures, and languages. Historically, diverse, and often conflicting, instructional approaches have

been implemented to meet the educational needs of children from these immigrant families. These have ranged from “English-only” approaches, which prohibit school use of the students’ native language (L1), to bilingual education approaches that seek to help students maintain and develop their L1 while acquiring English. For instance, at the turn of the 20th century, approximately 600,000 elementary school children—4% of all American students in the elementary grades—received either part or all of their instruction in the German language (National Association for Bilingual Education, 1998). California’s Proposition 227, on the other hand, legitimizes an English-only immersion approach, requiring that limited English proficient students be placed in a program in which all instruction is in English (Rossell, 2002). Since English language learners (ELLs) comprise an increasing proportion of the nation’s multicultural classrooms, it is crucial for educators to become more knowledgeable about the effects of L1 instruction on the acquisition of literacy skills in a second language (L2) and to use this knowledge to further the academic development of these students.

In contrast to the English-only paradigm (Rossell, 1988) advocated by opponents of bilingual education, the theoretical framework for two-way bilingual education programs is rooted in the transactional relationship between the L1 and L2. Nguyen, Shin, and Krashen (2001) assert that the instruction in the L1 is not detrimental to the development of English-language skills. In fact, it may even accelerate acquisition of the L2 and the development of academic skills in the L2, including reading, writing, and all the other skills necessary to succeed in school. A meta-analysis conducted by Greene (1997) indicates that the use of L1 instruction for ELLs is likely to improve student achievement, as measured by standardized tests in English.

As an explanation of this transactional relationship between the two languages, Cummins (1993) posed the interdependence hypothesis, which states that there is a transfer of knowledge, skills, and processes across languages, and that the development of L1 literacy skills facilitates the acquisition of academic skills in an L2. Students with well-developed literacy skills in their L1 are more likely to succeed in the acquisition of academic skills in an L2 than students whose L1 skills are not developed (Cummins, 1980). A fundamental tenet of Cummins’s interdependence hypothesis is that there is a common underlying language proficiency that determines an individual’s performance on cognitive and academic tasks.

## **Literature Review**

Unfortunately, the variety in existing bilingual programs has caused difficulty in interpreting research results in American classrooms. Indeed, the research literature identifies several types of programs that fall within the category of bilingual education but have different goals: (a) transitional or early-exit programs that provide L1 instruction in subject-matter studies while

the students are acquiring English; (b) late-exit programs that emphasize full bilingualism (i.e., the ability to speak, read, and write proficiently in two languages) and academic learning for ELLs; and (c) two-way bilingual education programs, which allow ELLs and fluent English speakers to learn English and another language. In contrast, the English-only submersion, or “sink or swim,” approach does not provide ELLs with any special help in overcoming language barriers that hamper students in learning content. While English-only programs may facilitate the acquisition of conversational English at a faster rate than bilingual programs, only transitional, late-exit, and two-way programs support the acquisition of academic English (Krashen, 1996).

However, not all research studies provide detailed information about the program(s) being studied (Porter, 1997). As a result, opponents of bilingual education label bilingual education programs a “failure” because “too many children have failed to become fluent in English” (“Bilingual Programs,” 1985, p. 1). These generalized statements rarely specify which type of bilingual program is the culprit.

Collier (1992) analyzed the results of several studies conducted in bilingual education and concluded:

If a study is conducted for enough years (five or six), not only do the bilingually schooled students outperform their [ELL] comparison group, making much greater gains, but they begin to reduce the gap between their performance on standardized tests, achieving as high or higher than 50 percent of the native speakers on a given test. (p. 193)

Thomas and Collier (1997) analyzed student performance based on the bilingual program in which the students were enrolled and found that kindergarten students who began school in the United States with no proficiency in English and were placed in a two-way, late-exit bilingual education program (which included both ELLs and fluent English speakers) outperformed the students in one-way, late-exit programs (those involving only ELLs), transitional programs, and programs with an English as a Second Language component. Moreover, according to student records collected from 1996 to 2001, students in two-way bilingual programs reached or surpassed the 50th percentile in both L1 and L2 in all subjects through the end of their schooling. This represents typical native English speakers’ performance on standardized tests (Thomas & Collier, 2002).

Ramírez (1992) compared ELLs’ performance on three treatments (English-only, transitional, and late-exit programs), in which the differentiating characteristic was the amount of instruction in the students’ L1. In this longitudinal study, Ramírez found that although many people believe that instruction time in the L1 is time lost to learning English, most students remain in English-only and transitional programs longer than these people expect; students do not exit English-only programs in 1 year. Ramírez concluded that “providing LEP students with substantial instruction in their primary language

does not interfere with or delay their acquisition of English language skills, but helps them ‘catch-up’ to their English-speaking peers in English language arts, reading, and math” (p. 1).

### **Purpose of Study**

The purpose of the present study is to investigate the effect of a two-way bilingual education program on the language development of English for Speakers of Other Languages (ESOL) instruction (Level 3 or 4) kindergarten and first-graders, as well as students who are proficient English speakers (ESOL Level 5 or native English speakers). ELLs, as well as proficient English speakers, were included in an Extended Foreign Language (EFL) program that sought to develop and maintain students’ oral and literacy skills in two languages: English and Spanish. Literacy was defined to include both reading and writing. In this study, the students in the experimental group all came from the EFL program and were instructed in English approximately 70% of the time and in Spanish approximately 30% of the time. The students in the control group, who were in mainstream classes in the same school, received instruction in English about 90% of the time and in Spanish (language arts) about 10% of the time. (The most significant difference between the groups was the amount of instruction time in Spanish, using the same curriculum.) Both the experimental and control groups consisted of ELLs and fluent English speakers. The academic performance of the students in the two groups was compared via administration of pretests and posttests, most of which had been developed by the district.

### **Methodology**

#### **Setting**

The study was conducted in an elementary school, in a predominately Spanish-speaking school district in the South. Immigration has a great impact on the school district, and on this school in particular, as a considerable proportion of its students come from immigrant families. Although 90% of the school’s students are native Spanish speakers, only 34% of the students are currently classified as ELLs. Approximately 33% of the school population consists of former ELLs who exited the ESOL program and are currently in mainstream classes. The socioeconomic background of the school population is diverse; 56% of the students receive free or reduced lunch, and most of the remaining students are middle class. During the 1999–2000 school year, reading scores were above the district average; however, a discrepancy existed between the achievement of ELLs and fluent English speakers, with the latter scoring higher than the former.

## Programs

Two programs within the same school were studied: the EFL program and the mainstream program. In the EFL program, a two-way bilingual program, English and Spanish were used as mediums of instruction. The goal was for both groups of students (ELLs and proficient English speakers) to become bilingual and biliterate (the number of students in each category is shown in Table 1). In this study, a different model was adopted than the typical approaches, which use 10% English instruction and 90% Spanish, or 50% of each language. As previously stated, students in the EFL group received instruction in English 70% of the day and Spanish 30% of the time. On the other hand, the ESOL model, which was provided to ELLs in the mainstream classes, sought to teach English-only except for the 10% of the school day spent on Spanish language arts and to mainstream the students into the regular program as quickly as possible.

## Participants

Participants in the study were 87 kindergarten students (43 males and 44 females) and 128 first-grade students (75 males and 53 females). Participants were classified as ESOL Level 3, 4, or 5 (Level 5 consisted of those who had exited the ESOL program and were fluent English speakers). Native English speakers also participated. In this school district, students classified as ESOL Level 3 and 4 receive special instruction in ESOL as part of their language arts program. Students classified as Level 5 no longer participate in ESOL, but their academic performance is monitored for 2 years after they exit the program. Students classified as gifted or learning disabled were excluded from the analyses, as they did not participate fully in the EFL or mainstream programs but were “pulled out” to receive special instruction during part of the school day.

This study used two experimental classes (EFL) in kindergarten and two in the first grade, as well as two control classes (mainstream) in each grade. There were preexisting differences between the experimental and control groups, in terms of language proficiency, for both the kindergarten and first-grade students. Table 1 presents the demographic and language proficiency composition of the two groups. A greater percentage of students were classified as ESOL in the experimental group than in the control group (74% vs. 15% in kindergarten, and 67% vs. 55% in the first grade). Furthermore, the kindergarten experimental group included a greater percentage (57%) of children on free or reduced lunch than the control group (44%). The percentage of students on free or reduced lunch has been shown to be a reliable indicator of socioeconomic status (SES), as well as a good predictor of achievement (Teddlie & Stringfield, 1993). Informal assessments in Spanish indicated no significant difference between the groups in reading, writing, and listening ability, although the students in the experimental group tended to score lower on these assessments than the students in the control group.

Table 1  
*Demographic and Other Characteristics of Experimental and Control Groups*

	Kindergarteners				First graders			
	Experimental group		Control group		Treatment group		Control group	
	<i>n</i>	Percentage	<i>n</i>	Percentage	<i>n</i>	Percentage	<i>n</i>	Percentage
Gender								
Male	22	48	21	51	33	58	42	59
Female	24	52	20	49	24	42	29	41
Lunch status								
Free or reduced	26	57	18	44	27	47	32	45
Full price	20	43	23	56	30	53	39	55
English for Speakers of Other Languages (ESOL) enrollment								
ESOL	34	74	6	15	38	67	39	55
Non-ESOL	12	56	35	85	19	33	32	45
<b>Total</b>	<b>46</b>	<b>100</b>	<b>41</b>	<b>100</b>	<b>57</b>	<b>100</b>	<b>71</b>	<b>100</b>

### Instruments

This study employed two sets of instruments to measure students' literacy development. One set measured the literacy growth of kindergarten students and was given to both the control and experimental groups. The other set measured the literacy development of first-grade students. Most of these assessments (described below), which were used to measure literacy in English, had been developed by the district and were widely used to measure the literacy goals and objectives delineated in the district's language arts curriculum. All assessments, except the Scholastic Reading Inventory (SRI) given to first graders, were administered at the beginning and at the end of the school year.

### *Kindergarten assessments*

The Kindergarten Assessment Guide includes a diagnostic survey of alphabet knowledge: upper-case and lower-case letters, letter production, and letter sounds. The guide also provides a rubric to classify student writing. Students are given a piece of paper and asked to draw a picture and write about it; the rubric is used to evaluate their ability to write. Students were also asked to read sight words from the list of High Frequency Words From Children's Literature.

### *First-grade assessments*

The "Emergent Reader Screening Assessment," used to measure the literacy skills of first-grade students, consists of four subtests. Of these, three provided information that was included in this analysis. The "Alphabet Knowledge" subtest measures students' ability to recognize upper-case and lower-case letters, and write and identify the sounds of the letters of the English alphabet. The "Phonemic Awareness" subtest requires that students match letter sounds to corresponding pictures, demonstrate their understanding of the graphophonic relationships in a spelling exercise, and identify the beginning and ending sounds in selected words. The third subtest requires the teacher to conduct a running record of students' ability to read. Given a booklet, each student is required to read while the teacher uses a sheet of paper, containing the full text that the student is asked to read, to take notes on the words read correctly, the miscues and self-corrections the students make as they read, and the overall fluency and comprehension that students demonstrate during their reading. A total score for the running record is calculated by adding the various subscores: the total of the points awarded for accuracy, self-corrections, use of cueing systems, fluency, and comprehension (with a maximum total of 20 points). The first-grade students were also asked to read words from the list of "High Frequency Words From Children's Literature." Furthermore, two writing samples, narrative and expository, were collected from each student. These prompts, which the district developed, were scored by two teachers using the district's primary grades rubric. Finally, the SRI was administered at the end of the school year. This test, which measures competency in reading comprehension, was administered as a part of the district's end-of-the-year evaluation. The SRI provides a Lexile score, a number that identifies the reading level of the child and represents his or her level of ability to read and comprehend passages (Scholastic, 2004).

### *Procedure*

The study used a two-group pretest–posttest design. The experimental group consisted of students participating in the EFL program, while the control group consisted of students in mainstream classes. Assignment of students to the experimental and control groups was not random. A letter was sent

home to parents informing them of the EFL program at the school. Interested parents filled out a form to register their children in the program. According to district recommendations, only students classified as ESOL Level 3, 4, or 5 and native English speakers could participate. The children whose parents filled out an EFL form but who were classified as ESOL Level 1 or 2 at the beginning of kindergarten were removed from the program, thus leaving vacancies in the EFL program. These vacancies were filled, at random, by calling the homes of children in ESOL Levels 3, 4, and 5, as well as native English speakers, to describe the program and offer the parents the opportunity to enroll their children.

Students participating in the EFL program were placed in one of two homerooms: the English language arts and social studies classroom, or the mathematics and Spanish language arts and science classroom. The students switched classes in the middle of the school day to receive instruction from the other teacher. These students received 2 hours of language arts in English, 30 minutes of independent reading time, 30 minutes of social studies in English, 1 hour of mathematics in English, 1 hour of language arts instruction in Spanish, and 30 minutes of science in Spanish.

Students classified as ESOL Level 3, 4, or 5 and native English speakers who did not participate in the EFL program were assigned to one of two self-contained classes. Students in the control group received all instruction in English, except for a weekly average of 2.5 hours of language arts in Spanish.

Once the groups were formed, the aforementioned pretests were administered at the beginning of the school year, and instruction in all classrooms reinforced the goals and objectives of the district's curriculum, as well as the benchmarks provided by the state standards. As the teachers worked in collaborative planning teams, the only difference in the official curriculum was the amount of time allowed for instruction in English and Spanish. District specialists and school site administrators monitored the EFL program. At the end of the school year, the posttests were administered, and the results were analyzed.

## **Results and Discussion**

As was previously discussed, a pretest–posttest control group design was used for this study. Since parents had the choice of selecting the type of program for their children, the assignment of children to the experimental and control groups was not random. As is evident in Table 1, there were slight differences between the experimental and control groups in terms of the proportion of students classified as ESOL and who received free or reduced lunch. This data indicates that students in the experimental group, especially kindergarteners, were at a disadvantage in terms of SES and in the proportion of ELLs.



## Kindergarten Findings

On the pretest scores of the kindergarten students, multivariate analysis of variance (MANOVA) with two factors (treatment group and SES level) and three dependent variables (alphabet knowledge, sight word mastery, and writing skill) indicated significant differences between the pretest scores of the experimental and control groups (Wilks Lambda = .28,  $F [3, 81] = 70.251$ ,  $p < .01$ ). Table 2 presents means and standard deviations. These initial differences were expected and pointed to the fact that the experimental group had a greater need for a special program. Indeed, research indicates that students entering kindergarten with developed phonemic awareness skills demonstrate significantly higher reading knowledge skills in subsequent school years (National Center for Education Statistics, 2002), thus suggesting the special need of the experimental group, who entered kindergarten at a disadvantage. Main effect of SES (free or reduced lunch vs. full-price lunch) on the combination of dependent variables was marginally significant (Wilks Lambda = .91,  $F [3, 81] = 2.66$ ,  $p = .05$ ). The interaction of group and SES was not significant ( $F [3, 81] = 0.46$ ,  $p = .71$ ).

As advised by Tabachnick and Fidell (2001), following the significant multivariate main effect of group on the combination of three dependent variables, univariate main effects of group on each single dependent variable were examined. In order to reduce Type I error due to multiple testing, a Bonferroni-adjusted  $\alpha$  (Tabachnick & Fidell, 2001) of .017 (.05/3) was used for each of the three univariate ANOVAs. The main effect of treatment on the writing scores was statistically significant ( $F [1, 83] = 178.09$ ,  $p < .01$ , eta-squared = .07). Group differences in the other two dependent variables were not statistically significant ( $F [1, 83] = 0.06$  for alphabet, and  $F [1, 83] = 1.41$  for sight words). When the main effect of SES on dependent variables was explored, the only significant effect pertained to alphabet ( $F [1, 83] = 7.65$ ,  $p < 0.01$ , eta-squared = .08). Main effect of SES on the other two variables was not significant ( $F [1, 83] = 1.26$  for sight words, and  $F [1, 83] = 0.31$  for writing).

On the posttest scores of the kindergarten students, multivariate analysis of variance with the three posttest scores as dependent variables revealed a main effect of group (Wilks Lambda = .79,  $F [3, 81] = 7.03$ ,  $p < .01$ ). Neither the main effect of SES, nor its interaction with treatment, was significant ( $F [3, 81] = 0.99$ , and  $F [3, 81] = 0.76$ , respectively). The mean scores of the two groups are presented in Table 2. Univariate tests of the group main effect, using a Bonferroni-adjusted  $\alpha$  of .017 (.05/3) indicated that there were no significant differences between the experimental and control groups in writing ( $F [1, 83] = 0.281$ ,  $p > .017$ ) or in sight words ( $F [1, 83] = 4.14$ ). However, the experimental group significantly lagged behind the control group in the alphabet test ( $F [1, 83] = 10.55$ ,  $p < .017$ , eta-squared = .11).

A possible explanation for the higher alphabet test scores of the control group might be that the number of proficient English speakers was greater in that group than in the experimental group ( $n = 35$ , as compared to 12). An examination of the means within the non-ELL subsample indicated that among the low-SES children, those in the control group outperformed those in the experimental group ( $M = 97.21$  and  $75.00$ ,  $SD = 23.00$  and  $2.97$ , respectively). However, since there were only three low-SES students in the experimental group, this difference should be interpreted cautiously. Among the higher SES students, the difference in performance between those in the experimental group and those in the control group was small ( $M = 95.43$  vs.  $93.56$ , respectively, with  $SD = 7.26$  and  $3.47$ ).

Table 2

*Mean and Standard Deviation of Kindergarten Pretest and Posttest Scores*

Measures	Experimental group		Control group		Extended Foreign Language main effect	Eta-squared
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
Pretest scores						
Alphabet	39.26	33.12	43.37	32.64	0.062	0.01
Sight Words	4.63	8.99	7.2	8.97	1.406	0.02
Writing	1	0	3.1	1.04	178.091*	0.68
Posttest scores						
Alphabet	86.35	17.38	95.85	5.98	10.547*	0.11
Sight Words	59.28	66.09	90.12	62.8	4.141	0.05
Writing	5.33	1.03	5.2	1.23	0.281	0.01

\*  $p < .017$  (Bonferroni-adjusted  $\alpha$  of  $.03/5$ ).

A 2x2 follow-up MANOVA within the English-proficient students indicated that there were significant main effects of treatment ( $F [3, 41] = 9.16$ ,  $p < .001$ , Wilks Lambda =  $.60$ ) and SES ( $F [3, 41] = 3.26$ ,  $p < .03$ , Wilks Lambda =  $.81$ ), as well as their interaction ( $F [3, 41] = 4.76$ ,  $p < .006$ , Wilks Lambda =  $.74$ ). The only significant univariate effect of the treatment factor (EFL vs. control group) was in alphabet knowledge ( $F [1, 43] = 19.00$ ,  $p <$

.001, eta-squared = .31). The proficient English speakers in the control group scored higher than the proficient English speakers in the EFL group ( $M = 96.14$  vs.  $88.92$ , with  $SD = 13.26$  and  $5.93$ , respectively). The univariate main effect of SES on alphabet knowledge, as well as the interaction of SES and treatment, were also significant ( $F [3, 41] = 9.16, p < .001$  and  $F [3, 41] = 4.76, p < .006$ , respectively).

### First-Grade Findings

On the pretest scores of the first-grade students, a 2x2 MANOVA with two factors (treatment group and SES level) indicated that the main effect of group was significant (Wilks Lambda = .78,  $F [6, 119] = 5.53, p < .01$ ). Neither SES nor its interaction with group was significant. Univariate tests, using a Bonferroni-adjusted  $\alpha$  of .0083 (.05/6; see Tabachnick & Fidell, 2001) indicated that there were differences between the experimental and control groups in four of the six tests (“Alphabet” and “Running Record” were the exceptions). The mean scores of the two groups are presented in Table 3. As shown, at the beginning of the year, the experimental group had lower scores in all six areas, although some differences were relatively small.

On the posttest scores of the first-grade students, MANOVA on the six posttest scores, as well as the SRI (percent of the items answered correctly), pointed to significant differences between the experimental and control groups (Wilks Lambda = .88,  $F [7, 118] = 2.32, p < .05$ ). Neither the main effect of SES nor its interaction with treatment group was significant ( $F [7, 118] = 0.921$  and  $F [7, 118] = 0.725$ , respectively). Following the significant multivariate main effect of treatment, univariate effects were examined, testing each at a Bonferroni-adjusted  $\alpha$  of .007 (.05/7). Results did not indicate group differences in any of the seven dependent variables.<sup>1</sup> In other words, the two groups of children did not differ from each other in any of the seven indicators of achievement.

### Conclusion

The main goal of the present study was to examine the effectiveness of a two-way bilingual program in reducing the achievement gap in English language arts between students with limited English proficiency and those who were relatively more proficient in English. The experimental program was implemented in kindergarten and first grade. Our hypothesis was that the achievement gap between the experimental and control groups would decrease or completely disappear as a result of the intervention. Our study found that after 1 academic year, there were no statistically significant gaps between the achievement scores of the experimental and control groups (this was true for both kindergarteners and first graders). The effects of SES on student achievement (i.e., the fact that the more children in the experimental group came from low-SES families) might have compounded the results for the kindergarten group, explaining the slight differences the study did find.

Table 3

*Mean and Standard Deviation of First-Grade Pretest and Posttest*

Measures	Experimental group		Control group		Extended Foreign Language main effect	Eta-squared
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
Pretest scores						
Alphabet	90.32	10.46	91.59	10.49	0.534	0.01
Phonics	61.65	19.76	71.55	22.35	7.173*	0.06
Running Record	1.3	0.6	1.32	0.79	0.047	0.01
Sight Words	18.09	18.8	56.73	77.41	12.917*	0.09
Narrative Writing	1.09	0.29	1.31	0.55	7.746*	0.06
Expository Writing	1.09	0.29	1.62	0.74	24.846*	0.17
Posttest scores						
Alphabet	96.89	2.78	96.49	9.43	0.163	0.01
Phonics	89.88	9.31	87.72	12.38	1.342	0.01
Running Record	2.82	0.83	3.03	1.04	1.417	0.01
Sight Words	208.6	128.38	278.27	163.47	6.65	0.05
Narrative Writing	2.18	0.98	2.51	1.01	3.539	0.03
Expository Writing	2.35	1.06	2.65	1	2.585	0.02
Scholastic Reading Inventory	38.38	17.21	44.44	25.95	1.981	0.02

\*  $p < .0083$  (Bonferroni-adjusted  $\alpha$  of .05/6).

The literature (e.g., Collier, 1992; Thomas & Collier, 1997; Weaver & Padron, 1999) has pointed to substantial gaps between the achievement of ELL and non-ELL students. In the present study, such a substantial gap was not present after 1 year of the two-way bilingual program intervention. The results from our study suggest that, although students in the experimental group had lower test scores at the beginning of the year, by the end of the year the achievement gap between the experimental and control groups had narrowed. This supports the potential usefulness of two-way bilingual programs in reducing the achievement gap between students with limited English proficiency and those whose English skills are more developed.

As previously discussed, the students in the control group received 90% of their instruction in English, with approximately 10% (2.5 hours a week) devoted to Spanish language arts. On the other hand, students in the experimental group received approximately 70% of their instruction in English, with Spanish used approximately 30% of the time (7.5 hours) for language arts and science. Thus, this study found that the additional time devoted to instruction in students' L1 did not hinder their academic progress in English language arts and the content areas. Furthermore, the study found that students in the experimental group, who began with lower test scores, scored comparably on the end-of-the-year tests after 1 year of being in the two-way bilingual program. One possible explanation for ELLs' ability to catch up to their peers is Cummins's interdependence hypothesis (1993). The time devoted to literacy instruction in the L1 (Spanish) facilitated the acquisition of the L2 (English). Nonetheless, future studies should look at the impact of instructional time in the L1 on student achievement. In this study, a mere 20% of additional instructional time devoted to the L1 helped partially "level the playing field" for ELL students. It could be argued that with more time devoted to L1 instruction, other programs could help students not only "catch up" to their peers but outperform them, as well.

Although the two-way bilingual program is the most plausible explanation for the narrowing of the achievement gap in the present study, there might be other possible causes that cannot be ruled out. It is theoretically possible, but not strongly plausible, that the two groups differed in their parental background in a manner that benefited the lower scoring experimental group. As was discussed, the experimental group came, generally, from lower SES families. It is highly conceivable that the parents of these children had low proficiency in English. Both of these factors (lower SES and lower English proficiency) are expected to disadvantage the experimental group of children, rather than being advantageous to them. It is conceivable that the impact of the intervention would have been even greater if these disadvantaging factors had not been present.

The present study examined the short-term effects of a two-way bilingual program at one school in the South. In order to better understand the impact of the EFL model, there is a need for new studies that follow up on students throughout their elementary school years (most two-way programs are K–6). Based on the growth trajectories in the present study, one may predict that the EFL students would, over time, demonstrate reading or writing performance on par with, or better than, the non-EFL students. Testing this prediction remains for future studies that would focus on the long-term effects of the model.

Finally, the results of the current study do not provide any insight regarding the impact of the EFL program on the L1 (Spanish) competency of the children. Since there is an interdependence between the L1 and L2 (Cummins, 1993), it is theoretically expected that the EFL group would be equally or more proficient in their L1 as well. That is, the development of literacy skills in the L1 (Spanish) facilitates acquisition of the L2 (English) and, in turn, literacy skills in the L2 have a positive impact on literacy skills in the L1. Further research is needed to confirm this prediction.

This study investigated the effects of a two-way bilingual program on the literacy development of ELLs, former ELLs, and native English speakers by comparing the performance of these students to that of a group of ELLs and fluent English speakers who attended the mainstream program at that school. Although there were significant differences between the two groups that favored the mainstream group at the onset of the study, after 1 academic year, it was found that the gap between the two groups had narrowed, and there were no significant differences between the two groups. Future studies are needed to identify effective bilingual models and learn more about the potential benefits of two-way bilingual programs.

## References

- Bilingual programs have “gone astray”—Education secretary eyes change in policy. (1985, September 26). *Seattle Times*. Retrieved January 21, 2004, from <http://proquest.umi.com/pqdweb?index=0&did=000000052905385&SrchMode=1&sid=1&Fmt=3&VInst=PROD&VType=PQD&TS=1074538679&clientId=30372> [ProQuest database].
- Collier, V. P. (1992). A synthesis of studies examining long-term language minority student data on academic achievement. *Bilingual Research Journal*, 16(1&2), 187–212.
- Cummins, J. (1980). The construct of language proficiency in bilingual education. In J. E. Alatis (Ed.), *Current issues in bilingual education* (pp. 81–103). Washington, DC: Georgetown University Press.

- Cummins, J. (1993). Bilingualism and second language learning. *Annual Review of Applied Linguistics*, 13, 51–70.
- Greene, J. P. (1997). A meta-analysis of the Rossell and Baker review of bilingual education research. *Bilingual Research Journal*, 21(2&3), 103–22.
- Krashen, S. D. (1996). *Under attack: The case against bilingual education*. Culver City, CA: Language Education Associates.
- National Association for Bilingual Education. (1998, Spring). History of bilingual education. *Rethinking Schools*, 12(3). Retrieved February 20, 2004, from [http://www.rethinkingschools.org/archive/12\\_03/langhst.shtml](http://www.rethinkingschools.org/archive/12_03/langhst.shtml)
- National Center for Education Statistics. (2002). *Children's reading and mathematics achievement in kindergarten and first grade*. Retrieved January 13, 2004, from <http://nces.ed.gov/pubs2002/kindergarten>
- Nguyen, A., Shin, F., & Krashen, S. (2001). Development of the first language is not a barrier to second language acquisition: Evidence from Vietnamese immigrants to the United States. *International Journal of Bilingual Education and Bilingualism*, 4(3), 159–164.
- Porter, R. P. (1997). The politics of bilingual education. *Society*, 34(6), 31–39.
- Ramírez, J. D. (1992). Executive summary of the final report: Longitudinal study of structured English immersion strategy, early-exit and late-exit transitional bilingual education program for language-minority children. *Bilingual Research Journal*, 16(1&2), 1–62.
- Rossell, C. H. (1988). The problem with bilingual research: A critique of the Walsh & Carballo study of bilingual education projects. *Equity and Excellence*, 23(4), 25–29.
- Rossell, C. H. (2002). *Dismantling bilingual education implementing English immersion: The California initiative*. San Francisco: Public Policy Institute of California. (ERIC Document Reproduction Service No. ED467043)
- Scholastic. (2004). *All about Lexiles*. Retrieved January 19, 2004, from <http://teacher.scholastic.com/sri/about.htm>
- Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics*. Boston: Allyn & Bacon.
- Teddlie, C., & Stringfield, S. (1993). *Schools make a difference: Lessons learned from a 10-year study of school effects*. New York: Teachers College Press.

- Thomas, W. P., & Collier, V. (1997). *School effectiveness for language minority students* (National Clearinghouse for Bilingual Education Resource Collection Series, No. 9). Washington, DC: Office of Bilingual Education and Minority Languages Affairs. (ERIC Document Reproduction Service No. ED 436087)
- Thomas, W. P., & Collier, V. P. (2002). *A national study of school effectiveness for language minority students' long-term academic achievement*. Washington, DC: Center for Research on Education, Diversity & Excellence. (ERIC Document Reproduction Service No. ED475048)
- Weaver, L., & Padron, Y. (1999). Language of instruction and its impact on educational access and outcomes. In A. Tashakkori, & S. H. Ochoa (Eds.), *Education of Hispanics in the United States: Politics, policies, and outcomes* (pp. 75–92). New York: AMS Press.

### **Acknowledgments**

We would like to express our appreciation for the thoughtful suggestions and comments of the reviewers and editors who have helped us improve the quality of this paper.

### **Endnote**

<sup>1</sup> To test the possible impact of an unequal representation of non-ELL students in the experimental ( $n = 19$ ) and control ( $n = 35$ ) groups, a separate follow-up MANOVA was performed within the non-ELL students only. There were no significant differences in academic performance of the experimental and control groups ( $F [7, 41] = 0.69$ ,  $p = .68$ ). Neither the SES main effect nor its interaction with treatment group was significant ( $F [7, 41] = 0.98$ ,  $p = .46$  and  $F [7, 41] = 0.56$ ,  $p = .78$ , respectively).