# Differential Outcomes of Two Bilingual Education Programs on English Language Learners

María G. López and Abbas Tashakkori
Florida International University

## Abstract

This study investigated the effects of two types of bilingual programs (two-way and transitional) on the academic performance and attitudes of fifth-grade students who entered kindergarten or first grade with different levels of English proficiency. A mixed methods design with both quantitative and qualitative data collection and analyses phases was employed. Quantitative data analyses indicated no significant differences in standardized measures of English achievement, although significant differences were found in other measures, including measures of oral language acquisition in English, Spanish-reading ability, students' attitudes, and perceived levels of proficiency in English and Spanish. Qualitative data analysis indicated that the students in two-way bilingual education programs were more likely to express positive attitudes towards bilingualism. Based on the mixed data, it is concluded that despite some similarity in the effects, each of the bilingual programs also has unique effects. Policy decisions should be made on the basis of relative importance, value, and the costs of these unique advantages and disadvantages.

## Introduction

Bilingual education is one of the most controversial topics in the field of education. In simplest terms, bilingual education, whether transitional or maintenance, is an instructional approach that uses the child's native language (L1) to make instruction meaningful. The controversy about bilingual education centers around the role of L1 instruction: Should English language learners (ELLs) receive instruction in English-only? Should they receive instruction in their L1 until they are able to comprehend English? Or should the schools

continue to develop the ELLs' L1 skills even after they have become proficient English speakers? Another question relates to the selection of students for types of bilingual programs: Which students should enroll in which programs?

In contrast to the "sink or swim" ideology advocated by opponents of bilingual education who believe that the students' L1 interferes with second-language (L2) acquisition and propose that ELLs be taught exclusively in English, the theoretical framework for two-way bilingual education (TWBE) programs is rooted in the interdependent relationship between the L1 and L2. Nguyen, Shin, and Krashen (2001) assert that the use of the L1 is not detrimental to the development of spoken English. In fact, it may even accelerate L2 acquisition and the development of academic skills in the L2. For instance, in a 2-year study, Carlisle and Beeman (2000) found that, on English academic assessments, children taught in Spanish did as well as the children taught in English, suggesting that instruction in L1 does not hinder L2 proficiency. Additionally, the students taught in Spanish demonstrated superior Spanish composition and reading comprehension skills.

Consequently, TWBE programs seek to facilitate the development of L2 skills while maintaining and enhancing the L1 skills and cultural integrity of students from both the minority- and majority-language groups. These programs attribute important roles to L1 and L2 respectively, as well as to the relationship between the two. Advocates of this model adhere to the paradigm influenced by Cummins' (1993) interdependence hypothesis, which purports that (a) there is a transfer of knowledge, skills, and processes across languages, (b) the development of L1 literacy skills facilitates the acquisition of academic skills in the L2, and, hence, (c) proficiency in L2 is a function of the level of L1 proficiency at the time when instruction in L2 begins.

Unfortunately, variety in program implementation has caused difficulty in interpreting research results in U.S. classrooms; however, the main thrust of studies to date (Collier, 1992; Ramírez, 1992; Thomas & Collier, 1997, 2002) provides evidence of the efficacy of TWBE programs. The differences between TWBE and other bilingual education programs are noted in the positive outcomes of academic development in language arts, reading, and mathematics, as demonstrated by the students' performance on standardized and non-standardized tests. For instance, López and Tashakkori (2004a, 2004b) found that TWBE reduced the achievement gap between ELLs and native English speakers. Similarly, Christian, Howard, and Loeb (2000) concluded that most of the studies based on TWBE indicate that students participating in those programs tend to perform as well as, or better than comparison groups enrolled in English-only or transitional bilingual education (TBE) programs. Christian et al. asserted that TWBE programs provide a sound basis for academic excellence for all students, as they facilitate the development of two languages.

Similarly, de Jong (2002) concluded that TWBE is an effective design for both native English-speaking and native Spanish-speaking students, although

by fifth grade, the ELLs in the TWBE program were still performing slightly below grade level on measures of English reading. When compared with other ELLs in the district and the state, however, the TWBE native Spanish speakers performed at higher levels. Additionally, in a study comparing two groups of Mexican American students, Curiel, Rosenthal, and Richek (1986) concluded that the longer students participated in bilingual programs, the less likely they were to drop out of school and the more likely they were to get better grades in school.

Given the high numbers of ELLs experiencing difficulties in school (Snow, Burns, & Griffin, 1998; Tashakkori, Ochoa, & Kemp, 1999), it is crucial that the effectiveness of the different models of bilingual education be studied—especially in times of reduced funding. That is, decisions regarding the curriculum imparted to language-minority students should be based on the careful examination of research findings and a consideration of different variables and conditions. Nonetheless, current policy in some schools dictates that children who have higher levels of English proficiency can participate in TWBE, while children with lower levels of English proficiency are discouraged from participating in programs that seek to develop literacy skills in both languages. This practice, which contradicts research findings (Hakuta, 1990; Ramírez, 1992; Thomas & Collier, 1997, 2002), suggests that TWBE and TBE programs have different results, depending on the students' level of English proficiency upon entering the program. An understanding of the possible differential effects of these two programs is critical to the decision-making progress influencing curricula and language policy.

The purpose of this study was to compare these two different approaches—TWBE (maintenance instruction in the L1) and TBE (transitional instruction in the L1)—among children with different L2 proficiency levels. In specific, the goal was to compare the academic performance and attitudes of fifth-grade students who had been enrolled in these two programs since kindergarten or first grade. Furthermore, this study attempted to determine whether differences between the two programs were related to the initial level of competency in English.

## Method

### Setting

This mixed method[1] (Tashakkori & Teddlie, 2003), causal-comparative study was conducted over a period of 1 year and involved six schools in a large school district in the southeastern part of the United States. Three of the six Bilingual School Organization (BISO)[2] schools in the district were purposefully selected for the study. These three schools were "matched," for the purposes of the study, to three other schools, which were similar to the

BISO schools in demographic characteristics but that employed a TBE model (i.e., the transitional English for speakers of other languages [ESOL] program whether in a self-contained or a pull-out setting).

## Programs

Two types of programs (i.e., TWBE and TBE) were studied; however, regardless of the program in which they enrolled, all students classified as ELLs received ESOL instruction. Those who were classified as *beginning* (level 1) or *intermediate* (level 2) received ESOL instruction and home language assistance in the content areas. Students classified as *advanced* (level 3) and *superior* (level 4), on the other hand, received 1 hour of ESOL instruction and 1 hour of regular language arts instruction. In TBE, no home language assistance was provided once the students reached this level of proficiency, since it was assumed that the children had acquired a certain level of proficiency in the L2 (English). TWBE programs continued to provide instruction in the native language to both ELLs and native English speakers throughout the elementary school years.

In a TBE program (the comparison group), once ELLs became somewhat proficient in English (i.e., once they attained a level 3 status), they no longer received content area instruction in their native language, although they had the opportunity to enroll in Spanish-for-speakers classes that consisted of 2.5 hours of instruction in Spanish-language arts every week. This comprised less than 10% of instructional time.

The students enrolled in a TWBE program (the treatment group), on the other hand, continued to receive content area and Spanish language arts instruction on a daily basis, even after they became proficient English speakers. Following the 60/40 model, the programs studied allotted 60% of instructional time to English and 40% to Spanish.

## Participants

A mixed multi-stage sampling strategy (Kemper, Stringfield, & Teddlie, 2003), consisting of both probability and purposive sampling, was utilized as described below. In the first stage, a purposive sample of six schools was selected. In the next stage of the sampling process, 553 fifth-grade students from the six identified schools were selected. Since the intent of the study was to determine the long-term effect of the two programs on the students' academic performance and attitudes, only students who had participated in the program at their school since kindergarten or first grade were considered.

Parent permission letters were sent home to the parents of all of the students identified as possible participants. The students who returned the signed permission slips (62.2% of the total) comprised the next stage of the sampling technique. As a validity check, the group that returned the permission

slips was compared with the group that did not participate on a variety of variables (e.g., gender, socioeconomic status [SES], and proficiency in English, learning exceptionality). Chi-square tests did not reveal any significant differences between the groups on any of these variables. A stratified non-proportional random subsample of 32 students was then selected from this group. These students completed the *Evaluación del Desarrollo de la Lectura* (*EDL*) (Ruiz & Cuesta, 2000) and interview. Special permission slips for the *EDL* and interview were randomly mixed with the ones that only requested permission for the questionnaire. Five to six students from each school were selected.

The absence of random assignment in an educational study poses threats to the internal validity of the conclusions. In order to overcome possible threats to inference quality related to differential selection, the TWBE schools were closely matched to TBE schools. The students included in the study were similar in terms of age and ethnicity. Chi-square tests indicated no significant differences between the TWBE and TBE groups in terms of English-language proficiency at entry level, SES, student exceptionality and number of retentions. Table 1 includes the demographic characteristics of the treatment and comparison groups. Lack of significant differences between the demographic characteristics of the two groups suggests that there were no significant differences that would cause one group to perform differently on measures of academic performance and/or attitudes.

## Variables and Their Measurements

### *Independent variables*

Two independent variables were considered—type of program (TWBE or TBE, as described above) and ESOL entry level at kindergarten (1, 2, 3, 4, or non-ESOL). Classification in these five groups had been done by the district 5 years earlier, when the children were in kindergarten. The district-developed Oral Language Proficiency Scale-Revised (OLPS-R) was used to assess the students' level of English proficiency. The test included various pictures to elicit oral responses. According to the score attained in the interview, students were labeled as ESOL level 1 (*beginning*), 2 (*intermediate*), 3 (*advanced*), 4 (*superior*), or 5 (*independent or non-ESOL*). Students who received ESOL instruction were those in levels 1–4. According to district statistics, test–retest reliabilities for the OLPS-R range from .80 to .94, and concurrent validity with the IDEA Oral Language Proficiency Test ranges from .72 to .81.

### *Dependent variables*

The dependent variables include academic achievement, affective, and cognitive measures. Academic achievement in English, including reading comprehension, mathematics, and science skills, was measured by the Florida

Table 1

*Demographic and Other Characteristics of the Experimental and Control Groups*

| Characteristic | Group | | | |
|---|---|---|---|---|
| | TWBE | | TBE | |
| | *n* | % | *n* | % |
| Gender | | | | |
| Male | 93 | 45.4 | 70 | 50.4 |
| Female | 112 | 54.6 | 69 | 49.6 |
| SES | | | | |
| Free lunch | 93 | 45.4 | 80 | 57.6 |
| Reduced-price | 37 | 18 | 21 | 15.1 |
| Regular-price | 75 | 36.6 | 38 | 27.3 |
| English proficiency | | | | |
| Level 1 | 56 | 27.3 | 59 | 42.4 |
| Level 2 | 23 | 11.2 | 16 | 11.5 |
| Level 3 | 27 | 13.2 | 13 | 9.4 |
| Level 4 | 25 | 12.2 | 10 | 7.2 |
| Non-ESOL | 74 | 36.1 | 41 | 29.5 |
| Exceptionality | | | | |
| Learning disability | 12 | 5.9 | 9 | 6.5 |
| Gifted | 40 | 19.5 | 25 | 18 |
| Other health impaired | 5 | 2.4 | 3 | 2.1 |
| Retention in grade | | | | |
| Retained | 10 | 4.9 | 5 | 3.6 |
| Not retained | 195 | 95.1 | 134 | 96.4 |

Comprehensive Assessment Test (FCAT), which measures the State's Sunshine State Standards. Content validity has been ascertained by several committees of experts (Florida Department of Education, 2002). Analyses conducted to test the reliability of the FCAT indicated that the FCAT scores have reliabilities similar to those of other standardized tests (Florida Department of Education).

Various instruments were used to measure other variables. Spanish-reading skills were measured with the *EDL*, the Spanish version of the Developmental Reading Assessment, which is an individually administered diagnostic instrument designed to identify a student's strengths and weaknesses in relation to reading accuracy, comprehension, and fluency. The validity and reliability of the instrument have been reported as high (Weber, 2000; Williams, 1999).

Attitudes towards bilingualism were measured via an interview and Likert-type questionnaire. The Likert-type scale consisted of eight items measuring the students' attitudes and values. They included the following: (a) bilingualism and intelligence, (b) bilingualism and the job market, (c) bilingualism and vocabulary development, (d) bilingualism and the school, (e) the bilingual program, and (f) bilingualism and the transfer of reading skills. Content validity was ascertained by teachers and students. Items included on the questionnaire were examined by experienced teachers in the areas of L2 acquisition, bilingual education, and reading instruction. Items on the questionnaire were those that the teachers identified as effective indicators and measures of the construct. Additionally, a pilot study was conducted involving a small group of fifth-grade students. These students were asked to examine the questionnaire to make sure that both the instructions and the items were comprehensible to fifth-grade students. Item-total correlations ranged from .33 to .61. Additionally, an interview, administered in Spanish, was utilized to gather information about the students' experience with their bilingual program and their facility with the L1.

The questionnaire also included items measuring the participants' self-reported level of proficiency in English and Spanish. One subsection, consisting of four items, asked the students to self-report on their level of proficiency in English, and a second subsection, also consisting of four items, asked the students to self-report on their proficiency in Spanish. The content of the items was taken from the National Education Longitudinal Study 88 Eighth Grade Questionnaire (United States Department of Education, 1988), although the items were modified to render them more appropriate to their intended audience.

## Procedures

As described above, three BISO schools that employ the TWBE model were selected for the study. These three schools were matched to three non-BISO schools that provided TBE to Spanish-speaking ELLs. Characteristics considered for the match included ethnic composition, percent of students receiving free and reduced lunch, percent of students identified as ELLs, and where possible, school size. Archival data such as demographics, ESOL information, and standardized test scores were collected at all sites.

Data collection included the administration of the questionnaire to all of the 344 students who returned the permission letter, as well as the Spanish *EDL* and follow-up interview to the randomly selected subsample of 32 students. All students received the same instructions and were provided with the same level of support during the administration. The *EDL* was administered as instructed in the accompanying Teacher's Manual.

Data analyses included both statistical and content analysis. Quantitative data were analyzed using statistical procedures, including several multivariate analyses of variance (MANOVA) and analyses of variance (ANOVA). Qualitative data were analyzed using both qualitative and quantitative techniques. Transcripts of the interviews were coded and analyzed for recurrent themes (Rubin & Rubin, 1995).

## Results

As Porter (1998) suggests, bilingual education means different things to different people. A recurrent criticism of research in bilingual education is that the program being studied is not described and, as a result, many programs that are labeled "bilingual" are simply classes of ELLs, homogenously grouped, for instruction in English. In order to avoid this problem, observations of eight classrooms were conducted in the TWBE schools. Each lasted an hour and was conducted during the Spanish language arts block. The types of interactions between teacher and students were recorded based on the language used (Buchanan & Tashakkori, 1998). During the Spanish language arts block, most (approximately 98%) of the teacher–student and student–teacher interactions were in Spanish. The use of English in teacher–student interactions was mostly to clarify terms and for managerial purposes. All of the lessons observed were teacher directed and provided little opportunity for student–student interactions (e.g., cooperative learning activities). The students were able to respond to the questions posed by the teachers and seemed to have mastered the vocabulary expected of them. While there were differences in the actual performance of the students from class to class and from school to school, the overall impression of all TWBE programs was that the students had been exposed to Spanish over prolonged periods of time.

### Findings From Quantitative Data Analyses

Quantitative data measuring academic and attitudinal variables were collected from the FCAT, *EDL*, and language questionnaire. Qualitative data from the interviews were quantified. Findings are presented by dependent variable.

*Academic achievement in English*

A MANOVA with two factors (i.e., type of program and ESOL level) and three dependent variables (i.e., fifth-grade FCAT reading, FCAT mathematics, and FCAT science) indicated that the effect of the type of program was not statistically significant, $F(3, 537) = 1.70$, $p = .17$. In other words, students who participated in TWBE did not score significantly differently on academic measures in English (i.e., reading comprehension, mathematics, and science) than the students who participated in TBE. Table 2 presents the means and standard deviations of the fifth-grade scores.

MANOVA, on the other hand, indicated that the main effect of ESOL entry level was statistically significant, $F(12, 1421) = 6.51$, $p < .001$, Wilks' Lambda = .87. Students who entered kindergarten or first grade with lower levels of proficiency in English scored lower than those more proficient in English, who, in turn, scored lower than the proficient English speakers. Because a significant main effect of ESOL entry level was found, univariate tests were conducted at a Bonferroni-type adjusted $\alpha$ level (Tabachnick & Fidell, 1989) of .17 (.05/3). All three univariate tests of main effects (i.e., FCAT reading, FCAT math, and FCAT science) were significant, $F(4, 539) = 16.35$, $p < .001$, $\eta^2 = .108$, $F(4, 539) = 8.12$, $p < .001$, $\eta^2 = .057$, and $F(4, 539) = 13.20$, $p < .001$, $\eta^2 = .089$, respectively. Interaction of the two factors (type of program and ESOL entry level) was not significant.

*Oral language development*

In order to determine whether a relationship exists between program participation and the amount of time needed to learn English, a univariate ANOVA was performed involving the same two factors but with the number of semesters enrolled in the ESOL program as a dependent variable. The main effect of type of program was statistically significant, $F(1, 356) = 5.72$, $p = .017$, $\eta^2 = .016$. Students in the TWBE programs spent an average of four semesters (2 school years) learning English, as compared to an average of five semesters (2.5 school years) for students in the TBE programs. Thus, consistent with previous research findings (Ramírez, 1992), students in the TWBE programs tended to require less time to exit the ESOL program and attain proficiency in English than their peers enrolled in TBE programs. Although this refutes the belief that time spent on the native language is time lost to English, it should be interpreted with caution given the small effect size ($\eta^2 < .02$). Type of program only accounted for less than 2% of variability in program length (i.e., vast proportion of variability was explained by other factors). However, this small magnitude of effect is still a strong confirmation for the conclusion that TWBE did not increase the amount of time required to become proficient in English. Mean number of semesters in ESOL and standard deviations are included in Table 3.

Table 2

*Academic Achievement as Measured by the FCAT*

| ESOL level | Academic achievement in English | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | FCAT Reading | | FCAT Math | | FCAT Science | | Number of participants | |
| | TWBE | TBE | TWBE | TBE | TWBE | TBE | TWBE | TBE |
| | *M* (*SD*) | *M* (*SD*) | *M* (*SD*) | *M* (*SD*) | *M* (*SD*) | *M* (*SD*) | *n* | *n* |
| 1 | 278 (55.06) | 286 (51.74) | 333 (49.24) | 325 (37.85) | 273 (58.78) | 279 (45.73) | 76 | 95 |
| 2 | 297 (57.76) | 312 (45.58) | 339 (45.44) | 337 (38.63) | 295 (43.06) | 306 (41.95) | 38 | 31 |
| 3 | 298 (51.31) | 303 (46.16) | 325 (55.97) | 333 (32.84) | 284 (52.05) | 288 (54.91) | 39 | 31 |
| 4 | 315 (43.10) | 294 (54.90) | 347 (56.91) | 339 (41.77) | 303 (50.73) | 305 (34.17) | 31 | 20 |
| Non-ESOL | 328 (50.27) | 323 (49.88) | 362 (47.58) | 345 (38.79) | 320 (55.25) | 307 (43.78) | 106 | 82 |
| Overall | 305 (55.52) | 303 (52.20) | 345 (51.70) | 335 (38.68) | 298 (56.83) | 294 (46.55) | 290 | 259 |

ANOVA also revealed a main effect of ESOL entry level, $F(3, 356) = 32.87$, $p < .001$, $\eta^2 = .22$. As expected, students with higher levels of English proficiency required less time to exit the ESOL program than their less proficient peers. There also was a significant interaction effect between ESOL entry level and the amount of time required to exit the ESOL program, $F(3, 356) = 4.64$, $p < .01$, $\eta^2 = .04$. Students with lower levels of proficiency in English (i.e., those classified as ESOL levels 1 or 2) exited the ESOL program at a faster rate when enrolled in a TWBE program; whereas, more proficient ELLs (i.e., those classified as ESOL levels 3 or 4) required less time in the ESOL program when enrolled in a TBE program.

### *Spanish reading*

A MANOVA with two factors (i.e., type of program and ESOL entry level) and three dependent variables (i.e., *EDL* Spanish-reading accuracy, *EDL* Spanish-reading comprehension, and *EDL* Spanish-reading fluency) indicated that the main effect of type of program was statistically significant, $F(3, 22) = 5.69$, $p < .05$, Wilks' Lambda = .56. As more instructional time was spent in

Table 3

*Means and Standard Deviations of the Number of Semesters in the ESOL Program*

| ESOL level | Oral language acquisition | | | | | |
|---|---|---|---|---|---|---|
| | TWBE | | | TBE | | |
| | *M* | *SD* | *n* | *M* | *SD* | *n* |
| 1 | 4.82 | 2.23 | 76 | 6.03 | 1.87 | 97 |
| 2 | 3.16 | 1.1 | 38 | 4.52 | 1.71 | 31 |
| 3 | 3.62 | 1.82 | 39 | 3.56 | 1.66 | 32 |
| 4 | 3.39 | 1.38 | 31 | 2.9 | 1.65 | 20 |
| Overall | 3.98 | 1.95 | 184 | 4.98 | 2.15 | 180 |

Table 4

*Means and Standard Deviations of Fifth-Grade EDL*

| ESOL level | Reading in Spanish | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Accuracy | | Comprehension | | Fluency | | Number of participants | |
| | TWBE | TBE | TWBE | TBE | TWBE | TBE | TWBE | TBE |
| | *M* (*SD*) | *M* (*SD*) | *M* (*SD*) | *M* (*SD*) | *M* (*SD*) | *M* (*SD*) | *n* | *n* |
| 1 | 97.75 (1.26) | 97.25 (2.87) | 15.25 (6.24) | 14 (4.40) | 4.88 (0.85) | 4.75 (0.29) | 4 | 4 |
| 2 | 96.33 (1.53) | 66 (45.31) | 11.67 (2.08) | 7.75 (8.02) | 5 (1.00) | 2.5 (2.08) | 3 | 4 |
| 3 & 4 | 77.8 (43.49) | 71 (47.34) | 9.8 (7.60) | 7 (5.77) | 3.8 (2.17) | 2.75 (1.89) | 5 | 4 |
| Non-ESOL | 95 (3.56) | 66 (44.21) | 9.5 (3.00) | 7.75 (8.02) | 4.63 (0.48) | 2.13 (1.55) | 4 | 4 |
| Overall | 90.56 (24.24) | 75.06 (37.82) | 11.44 (5.60) | 9.13 (6.69) | 4.5 (1.35) | 3.03 (1.78) | 16 | 16 |

*Note*. Due to small cell sizes, ESOL levels 3 and 4 were combined into one category.

Spanish in the TWBE schools, students enrolled in these programs scored higher in measures of Spanish than their peers enrolled in the TBE programs at the comparison schools. Means and standard deviations are shown in Table 4. Because a significant main effect of program was found, univariate tests were conducted on each dependent variable. A Bonferroni-type adjustment was made for an inflated Type I error due to multiple ANOVAs (Tabachnick & Fidell, 1989). The $\alpha$ level was set to .017 (.05/3). With the adjusted $\alpha$ level, only reading fluency (as measured by the *EDL*) was found to be significant, $F(3, 24) = 8.20$, $p < .017, \eta^2 = .255$. On the other hand, MANOVA indicated that neither ESOL entry level nor its interaction with type of program was statistically significant, $F(9, 54) = 0.84$, $p = .58$ and $F(9, 54) = 0.84$, $p = .58$, respectively.

### *Attitudes towards bilingualism*

ANOVA with two factors (i.e., type of program and ESOL level at kindergarten or first grade) and the average of the responses to the items pertaining to attitude towards bilingualism as the dependent variable indicated that type of program was statistically significant, $F(1, 325) = 25.60$, $p < .001, \eta^2 = .073$. Students enrolled in TWBE programs had more positive attitudes towards bilingualism in English and Spanish. Means and standard deviations are presented in Table 5. On the other hand, the ANOVA indicated that neither the main effect of ESOL entry level nor its interaction with type of program was statistically significant, $F(4, 325) = 0.61$, $p = .66$ and $F(4, 325) = 0.28$, $p = .89$, respectively.

Table 6 includes an item-by-item presentation of student responses. For simplicity, responses marked as "strongly agree" and "agree" are combined into the "agree" category, and responses marked as "strongly disagree" and "disagree" are combined into the "disagree" category. Chi-square tests were conducted to compare the responses of the TWBE and TBE students. To compensate for the number of statistical tests, an $\alpha$ level of .006 (.05/8) was utilized for these tests. Chi-square tests indicated significant differences between the two groups in their responses to the second item, $\chi^2(3) = 24.41$, $p < .001$; fifth item, $\chi^2(3) = 15.33$, $p = .002$; sixth item, $\chi^2(3) = 13.51$, $p = .004$; seventh item, $\chi^2(3) = 13.44$, $p = .004$; and eighth item, $\chi^2(3) = 20.38$, $p < .001$. Students in the TWBE programs were more likely than the students in the TBE programs to think that a language other than English should be learned at school, that knowing Spanish is helpful when reading new words in English, that learning two languages helps them think better, that learning two languages is useful, and that they enjoy learning two languages in school.

### *Self-reported level of language proficiency*

A MANOVA with two factors (i.e., type of program and ESOL level at kindergarten or first grade) and two dependent variables (i.e., self-reported level of proficiency in English and in Spanish) indicated that the main effect of

Table 5

*Means and Standard Deviations of the Students' Attitudes Towards Bilingualism*

| ESOL level | Attitudes towards bilingualism | | | | | |
|---|---|---|---|---|---|---|
| | TWBE | | | TBE | | |
| | *M* | *SD* | *n* | *M* | *SD* | *n* |
| 1 | 3.51 | 0.33 | 54 | 3.33 | 0.31 | 57 |
| 2 | 3.51 | 0.25 | 23 | 3.22 | 0.39 | 16 |
| 3 | 3.51 | 0.36 | 27 | 3.25 | 0.38 | 12 |
| 4 | 3.53 | 0.33 | 24 | 3.36 | 0.42 | 10 |
| Non-ESOL | 3.48 | 0.31 | 73 | 3.25 | 0.43 | 39 |
| Overall | 3.5 | 0.31 | 201 | 3.29 | 0.37 | 134 |

*Note*. The mean score was obtained from a Likert-type questionnaire measuring attitudes towards bilingualism. The four response options ranged from "strongly disagree" (1) to "strongly agree" (4).

type of program was statistically significant, $F(2, 325) = 27.88, p < .001$, Wilks' Lambda = .85. Mean scores and standard deviations are shown in Table 7. Because a significant main effect of program was found, follow-up univariate tests were conducted at an adjusted $\alpha$ level of .025 (.05/2). These tests indicated that the effects of program on self-perceptions of proficiency in English and Spanish were significant, $F(1, 326) = 5.44$, $p = .02$, $\eta^2 = .016$ and $F(1, 326) = 13.62$, $p < .001$, $\eta^2 = .103$, respectively. Students in TBE programs tended to report higher levels of proficiency in English than their peers in TWBE programs, while the TWBE students tended to report higher levels of Spanish proficiency than the students enrolled in the TBE programs.

MANOVA also revealed a significant main effect of ESOL entry level, $F(8, 650) = 6.19$, $p < .001$, Wilks' Lambda = .86. Univariate tests were conducted at an adjusted $\alpha$ level of .025 (.05/2). These tests revealed significant differences in self-reported level of English proficiency and self-reported level of Spanish proficiency, $F(4, 326) = 2.88$, $p < .025$, $\eta^2 = .034$ and $F(4, 326) = 2.46$, $p < .001$, $\eta^2 = .077$, respectively. Now, in fifth grade, students who were more proficient in English at kindergarten or first-grade entry considered themselves more proficient in English than those who were less proficient at kindergarten or first-grade entry. On the other hand, those who were less proficient in English at kindergarten or first-grade entry tended to consider themselves

Table 6

*Frequency and Percentages of Student Responses to the Questionnaire Items Measuring Attitudes*

| **Questionnaire item** | **TWBE** | | **TBE** | |
|---|---|---|---|---|
| | Agree | Disagree | Agree | Disagree |
| | *n* (%) | *n* (%) | *n* (%) | *n* (%) |
| 1. Learning two languages makes me smarter. | 191 (97.0) | 6 (3.0) | 126 (94.7) | 7 (5.3) |
| 2. We should learn only English at school. | 3 (1.5) | 192 (98.5) | 16 (12.0) | 117 (88.0) |
| 3. I think that knowing English and Spanish will help me get a good job in the future. | 198 (99.0) | 2 (1.0) | 127 (96.2) | 5 (3.8) |
| 4. Learning Spanish is useless. | 4 (2.0) | 195 (98.0) | 7 (5.3) | 126 (94.7) |
| 5. Knowing Spanish helps me read new words in English. | 123 (61.8) | 76 (38.2) | 54 (40.9) | 78 (59.1) |
| 6. Learning two languages helps me think better. | 175 (89.3) | 21 (10.7) | 102 (77.9) | 29 (22.1) |
| 7. Learning two languages is a waste of time. | 4 (2.0) | 190 (98.0) | 5 (3.8) | 128 (96.2) |
| 8. I like learning two languages in school. | 194 (97.5) | 5 (2.5) | 115 (86.5) | 18 (13.5) |

*Note*. Responses indicating strong agreement and agreement are combined into the "agree" category, while responses indicating strong disagreement and disagreement are combined into the "disagree" category.

more proficient in Spanish than their peers. Interaction of type of program and ESOL entry level was not statistically significant, $F(8, 650) = 1.61$, $p = .12$, Wilks' Lambda = .96.

Table 7

*Means and Standard Deviations of the Students' Self-Reported Level of Language Proficiency*

| ESOL level | Student self-report of their language proficiency | | | | | |
|---|---|---|---|---|---|---|
| | English proficiency | | Spanish proficiency | | Number of participants | |
| | TWBE | TBE | TWBE | TBE | TWBE | TBE |
| | *M* (*SD*) | *M* (*SD*) | *M* (*SD*) | *M* (*SD*) | *n* | *n* |
| 1 | 3.69 (0.41) | 3.79 (0.31) | 3.62 (0.48) | 3.43 (0.62) | 55 | 57 |
| 2 | 3.73 (0.26) | 3.77 (0.32) | 3.38 (0.61) | 2.94 (0.76) | 23 | 16 |
| 3 | 3.78 (0.37) | 3.94 (0.11) | 3.44 (0.49) | 3.06 (0.69) | 27 | 12 |
| 4 | 3.81 (0.28) | 3.93 (0.17) | 3.63 (0.39) | 2.6 (0.52) | 24 | 10 |
| Non-ESOL | 3.82 (0.29) | 3.89 (0.23) | 3.35 (0.55) | 2.93 (0.86) | 72 | 40 |
| Overall | 3.77 (0.34) | 3.84 (0.27) | 3.47 (0.52) | 3.13 (0.76) | 201 | 135 |

*Note*. The four response options to these questions ranged from "not at all" (1) to "very well" (4).

## Findings From Qualitative Data Analysis

As mentioned above, structured interviews were conducted with a randomly selected subsample of students. In the TWBE programs, these interviews were conducted in Spanish (with the exception of one native English-speaking student who requested to respond in English). Four themes were identified from the responses of the students[3] in the TWBE programs as follows.

First, bilingualism is an asset for the future. The students were quick to explain that the ability to read and write, not just speak, two languages would be advantageous to them in the future. As Luis, a student enrolled in one of the TWBE programs, stated, "*El programa [bilingüe] es muy bueno porque nos ayuda en el futuro con trabajos bilingües, especialmente en lugares como Miami, donde hay muchos latinos, y en California . . .* (The [bilingual]

program is very good because it helps us in the future with bilingual jobs, especially in places like Miami, where there are many Latinos, and in California . . .)." Moreover, the students were able to imagine hypothetical instances in the future when knowing two languages would facilitate their ability to provide effective customer service to monolingual clients.

Second, bilingualism facilitates communication between groups. Most of the students believed that knowing two languages would allow them to translate for monolinguals and thus ease communication between individuals and groups. More personally, bilingual instruction provided the keys to the two different worlds that the child inhabits. Martha explained how her TWBE program has contributed to her life:

> *Yo, cuando era chiquita, yo no sabia mucho inglés. Entonces a mi no me gustaba, pero después yo lo aprendí y ahora yo puedo hablar con mis amigas en la escuela y puedo hablar con mi familia en la casa porque mi familia habla en español. Ellos no hablan en inglés mucho* [I, when I was little, I did not know a lot of English. Then, I did not like it, but later I learned it and now I can talk to my friends at school and I can talk to my family at home because my family speaks in Spanish. They do not speak a lot of English].

Third, bilingualism maintains the cultural heritage. As Luis explained, the bilingual program is good because "*nos ayuda mucho y aprendo mucho, de mi cultura, de mis padres, y todo* [it helps us a lot, and I learn a lot about my culture, about my parents and everything]." In the sense that it develops oral skills in the L1, the TWBE program allows the students to communicate with monolingual family members and thus maintains cultural ties.

Fourth, bilingualism is an academic aid. Most TWBE students mentioned that bilingualism helps them read in two languages, and they were able to explain the role of cognates. As described by Maria, "*Algunas palabras en español se pueden parecer al inglés y entonces si, por ejemplo, te dicen alguna palabra y tu sabes cual es el significado en inglés, puedes leer la palabra [en español]* (Some words in Spanish can look like [words in] English and then if, for example, they tell you a word and you know what it means in English, you can read the word [in Spanish])." They believed that this facilitates the reading process, as it helps them develop an enriched vocabulary. Furthermore, they believed that reading strategies learned in one language transfer to reading in the other language.

Overall, the students enrolled in the TWBE programs expressed positive attitudes towards the instruction that they received in English and Spanish. They were confident that bilingualism would be an asset to them in future.

Interviews were also conducted in the TBE programs. Most students enrolled in these schools preferred to conduct the interview in English. Some of the students in the TBE programs expressed beliefs similar to those held by

the students in the TWBE group. However, differing voices were heard in this group. In effect, the students in the TBE programs could be divided into two groups. One group was quite similar to the TWBE group. However, the second group had a different profile. The following four themes emerged from their interviews.

First, bilingualism is an asset for the future. Like their peers in the TWBE programs, all of the students in the TBE programs expressed the belief that knowing two languages would help them get a better job.

Second, lack of bilingualism impedes communication between groups. This group of students expressed frustration because they were not able to communicate effectively with Spanish speakers. Over the course of the years in the TBE program, their L2 became noticeably dominant. After 5 years of schooling, Magda admitted that "it is hard for me to understand Spanish." Like some of her peers, she lost the ability to communicate with others in her native language. Because of the language shift (Wong-Fillmore, 2000) that she experienced, she was no longer able to communicate with others in Spanish. In fact, this child would probably need others to translate for her when trying to communicate with Spanish-speaking monolinguals.

Third, lack of bilingualism hinders the maintenance of cultural heritage. With regret, Jorge admitted that "I don't really know how to talk that well" in Spanish. Another student commented that his lack of facility with his L1 made communication with his father and grandfather difficult, as these two members of this family only speak Spanish.

Fourth, bilingualism is not an academic aid. In fact, instruction in two languages is confusing for this group of students. Because Spanish is difficult for them to understand, the two languages are seen as very different from each other. Magda believes that the two languages are really different, and Mark expressed frustration: "They have different words and different meanings sometimes. Spanish words may look the same as English words but they mean different things." Cognates are confusing rather than useful, and the reading strategies are described as being different.

This second group of students enrolled in the TBE programs experienced a certain degree of frustration because they lost command of their L1.

## Discussion

It was found that there were no significant differences between children who participated in TWBE and those who participated in TBE on measures of academic achievement in English as measured by the reading, mathematics, and science portions of the FCAT. However, those who participated in TWBE acquired oral English at a faster rate and performed better on measures of reading in Spanish. Students in TWBE programs were satisfied with their

language proficiencies and viewed bilingualism as an asset. TBE programs, on the other hand, did not affect all students in the same manner. Instead, two views existed in those programs: those who, like their BISO peers, saw bilingualism as an asset and those who, having experienced a language shift, felt frustrated by their inability to communicate effectively in Spanish.

Results indicate that, regardless of the type of program, students who were most proficient upon entering kindergarten or first grade also scored the highest on measures of academic achievement in English 5 years later in the fifth grade. The type of program in which they enrolled made no differential impact on their achievement. The alarming conclusion that might be made on the basis of these findings is consistent with the predictions of Rossell (2002) and the Coleman Report (Berliner & Biddle, 1995): The type of program designed for ELLs does not remove their initial disadvantage. This conclusion, however, is contradicted by other aspects of the results, especially when the qualitative and quantitative results are integrated. Although there were still differences between the ELLs and the native English speakers on standardized measures and neither program removed the pre-existing achievement gap, one must consider that the ELLs had to learn a L2, and, in some cases, they did this while they developed literacy skills in their L1 and positive attitudes towards bilingualism, as demonstrated by the other measures included in this study. Therefore, the programs, especially TWBE, did have a positive academic effect on the students.

The results of this study are consistent with the conclusions reached by Lambert and Tucker (1973) who found no differences in the English achievement of the bilingually schooled children and those who were taught monolingually. Looking at a different population and social context, Ollers and Pearson (2002) reached similar conclusions. Similarly, Carlisle and Beeman (2000) found that the English-reading scores of students taught in English did not differ significantly from those of students taught in Spanish. However, the two groups did differ on measures of reading in Spanish, with the group taught in Spanish outperforming the other group.

Although the results of the current study indicate that type of program did not differentially impact students' English proficiency as measured by the FCAT, there was confirmation that L1 instruction did not impede or delay L2 acquisition. On the contrary, L1 instruction accelerated the rate of L2 acquisition in the L2, while facilitating the maintenance and development of literacy skills in the L1. This finding is consistent with previous research (Lucido, 2000; Ramírez, 1992). Using the length of time in bilingual education as an indicator of effectiveness, instruction in L1 seems to have a positive impact. Students who received more instruction in their L1 (i.e., those in the TWBE programs) learned English the fastest, especially if they entered kindergarten with little or no proficiency in English. On average, they were classified as proficient English speakers approximately one semester earlier than the control group.

One of the reasons why this study was conducted was to determine whether the practice of excluding students with low levels of English-language proficiency from TWBE programs could be justified by research. The findings do not support this practice. Students with low levels of English proficiency who were taught in the TWBE programs did not underperform when compared with those taught in TBE programs. To the contrary, they acquired oral language at a faster rate and were either at the same level as, or exceeded, the other group.

In a time of reduced funding and movements toward standardization, it is important to consider the efficacy of different programs available to ELLs. Decisions about which students can and which cannot participate in TWBE programs should be based on a careful analysis of all of the issues involved. The results of this study indicate that TWBE produces similar effects in terms of academic achievement in English as TBE. However, students enrolled in TWBE acquire oral language at a faster rate, are more proficient in Spanish, and hold more positive attitudes towards bilingualism.

The findings present an interesting question regarding the current policy of using standardized tests (in this case, the state-mandated FCAT) as the sole determinant of the academic progress of ELLs: How should their growth be measured and monitored? Tests of oral language development indicate that one program is better than another for ELLs at different entry levels. Analyses of standardized test scores, however, indicate that the programs produce similar results. After 5 years of participation, neither program completely reduced the achievement gap. Nevertheless, this is not true. The ELLs who entered kindergarten with low levels of English proficiency have made significant academic growth: They acquired a L2. FCAT scores by themselves fail to demonstrate this achievement. Therefore, there is a need to employ multiple measures so as to make evident the progress of ELLs.

Given the lack of statistical significance in the FCAT scores of both groups, the continuation of TWBE programs may be called into question. However, the opposite can be argued. It must be noted that, while the two programs produced similar competencies in English language and the content areas, they also produced a difference in their ability to read and comprehend a passage in their L1. In other words, without a negative impact on the students' academic performance in English, TWBE programs facilitated the development of literacy abilities in the students' L1. As a result of participating in a program that emphasized bilingualism and biliteracy, the students became more proficient readers in both languages. As the profile of the TWBE students suggests, these students felt comfortable using both languages and found that there was a transfer across languages that facilitated literacy tasks in both English and Spanish. Moreover, the native English speakers enrolled in the TWBE programs also benefited from the bilingual model.

Previous research in bilingual education has predominantly utilized quantitative methods. The current study suggests that stronger and more trustworthy inferences are reached when a combination of quantitative and qualitative methods are utilized and incorporated. The qualitative component of the study provided a different level and type of insight than otherwise would have been achieved from the quantitative results.

Nevertheless, there are several limitations to this study that suggest the need for future research. A longitudinal design would enhance the inference quality by documenting the progress of the students in both groups from entry into kindergarten to the end of fifth grade. Such a design would also provide information about the small number of students who exit the programs before the end of fifth grade. Using parent questionnaires to collect additional demographic and home data would allow for a better comparison of the entry level characteristics. Due to the fact that BISO schools in the district are designed exclusively in English and Spanish, the investigation only studied TWBE programs involving Spanish. There is a need for further studies among students who speak other languages.

## References


Berliner, D. C., & Biddle, B. J. (1995). *The manufactured crisis: Myths, fraud, and the attack on America's public schools*. Cambridge, MA: Perseus Books.

Buchanan, T., & Tashakkori, A. (1998). *Evaluation of the Louisiana K–3 reading and math initiative: 1997–1998 school year*. Baton Rouge, LA: Louisiana Department of Education, Division of School Standards, Accountability, and Assistance.

Carlisle, J. F., & Beeman, M. M. (2000). The effects of language of instruction on the reading and writing achievement of first-grade Hispanic children. *Scientific Studies of Reading, 4*(4), 331–353.

Christian, D., Howard, E. R., & Loeb, M. I. (2000). Bilingualism for all: Two-way immersion education in the United States. *Theory Into Practice, 39*(4), 258–266.

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. (1993). Bilingualism and second-language learning. *Annual Review of Applied Linguistics, 13*, 51–70.

Curiel, H., Rosenthal, J. A., & Richek, H. G. (1986). Impacts of bilingual education on secondary school grades, attendance, retentions, and dropout. *Hispanic Journal of Behavioral Sciences, 8*(4), 357–367.

de Jong, E. (2002). Effective bilingual education: From theory to academic achievement in a two-way bilingual program. *Bilingual Research Journal, 26*(1), 65–84.

Florida Department of Education. (2002). *Technical report: For operational test administrations of the 2000 Florida comprehensive assessment test.* Retrieved January 13, 2006, from http://firn.edu/doe/sas/fcat.htm

Hakuta, K. (1990). *Bilingualism and bilingual education: A research perspective.* NABE Focus: Occasional Papers in Bilingual Education 1. Retrieved February 16, 2006, from http://www.ncela.gwu.edu/pubs/focus/focus1.htm

Kemper, E. A., Stringfield, S., & Teddlie, C. (2003). Mixed methods sampling strategies in social science research. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 273–296). Thousand Oaks, CA: Sage Publications, Inc.

Lambert, W. E., & Tucker, G. R. (1973). The benefits of bilingualism. *Psychology Today, 7*(4), 89–93.

López, M. G., & Tashakkori, A. (2004a). Effects of a two-way bilingual program on the literacy development of students in kindergarten and first grade. *Bilingual Research Journal, 28*(1), 63–78.

López, M. G., & Tashakkori, A. (2004b). Narrowing the gap: Effects of a two-way bilingual education program on the literacy development of at-risk primary students. *Journal of Education for Students Placed at Risk, 9*(4), 325–336.

Lucido, F. (2000). The influence of bilingualism on English reading scores. *Reading Improvement, 37*(2), 87–91.

Nguyen, A., Shin, F., & Krashen, S. D. (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.

Oller, D. K., & Pearson, B. Z. (2002). Assessing the effects of bilingualism: A background. In D. K. Oller & R. E. Eilers (Eds.), *Language and literacy in bilingual children* (pp. 3–21). Tonawanda, NY: Multilingual Matters Ltd.

Porter, R. P. (1998). The case against bilingual education. *The Atlantic Monthly, 281*(5), 28–39.

Ramírez, J. D. (1992). Executive summary. *Bilingual Research Journal, 16* (1–2), 1–62.

Rossell, C. H. (2002). *Dismantling bilingual education implementing English immersion: The California initiative*. San Francisco, CA: Public Policy Institute of California. (ERIC Document Reproduction Service No. ED467043)

Rubin, H. J., & Rubin, I. S. (1995). *Qualitative interviewing: The art of hearing data*. Thousand Oaks, CA: Sage Publications, Inc.

Ruiz, O., & Cuesta, V. (2000). *Evaluación del desarrollo de la lectura: Guía de recursos para el maestro* [Evaluation of the development of readings: Resource guide for the teacher]. Parsippany, NJ: Celebration Press.

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

Tabachnick, B. G., & Fidell, L. S. (1989). *Using multivariate statistics* (2nd ed.). New York: Harper & Row.

Tashakkori, A., Ochoa, S. H., & Kemper, E. A. (1999). Pull and push factors in the educational attainment of Hispanics: Current realities and future directions. In A. Tashakkori, S. H. Ochoa, & E. A. Kemper (Eds.), *Education of Hispanics in the United States* (pp. 249–268). New York: AMS Press, Inc.

Tashakkori, A., & Teddlie, C. (Eds.). (2003). *Handbook of mixed methods in social and behavioral research*. Thousand Oaks, CA: Sage Publications.

Thomas, W. P., & Collier, V. P. (1997). *School effectiveness for language minority students*. Washington, DC: National Clearinghouse for Bilingual Education. (ERIC Document Reproduction Service No. ED436087)

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)

United States Department of Education. (1988). *National education longitudinal study of 1988*. Washington, DC: Center for Education Statistics.

Weber, W. A. (2000). *Developmental reading assessment and Evaluación del desarrollo de la lectura: A validation study*. Retrieved January 13, 2006, from http://www.pearsonlearning.com/correlation/rsp_srch_temp.cfm?program=dra

Williams, E. J. (1999). *Developmental reading assessment: Reliability study*. Retrieved January 13, 2006, from http://www.pearsonlearning.com/correlation/rsp_srch_temp.cfm?program=dra

Wong-Fillmore, L. (2000). Loss of family languages: Should educators be concerned? *Theory Into Practice, 39*(4), 203–210.

## Endnotes

[1] According to Tashakkori and Teddlie (2003), a mixed method study is one that includes both quantitative and qualitative data collection and analysis, conducted concurrently or sequentially.

[2] Bilingual School Organization (BISO) means that the school is organized such that the TWBE, or dual language, model is implemented in all classrooms. All of the students in these selected BISO schools received instruction in English language arts, Spanish language arts, and content in both languages.

[3] To protect the students' identities, pseudonyms are used in this section.