

Effects of a Transitional Bilingual Education Program: Findings, Issues, and Next Steps

Carolyn Huie Hofstetter
University of California, Berkeley

Abstract

This study examined the effectiveness of a transitional Spanish–English bilingual program, Academic Language Acquisition (ALA), in enhancing K–5 students’ English-language proficiency, as well as their English performance in academic subject areas, in comparison with the Structured English Immersion (SEI) process.¹ An existing reading program, Success for All (SFA), served as a confounding influence because it had similar goals for reading development and included English- and Spanish-language curricula. Given 4 years of enrollment in their respective programs, ALA and SEI students, regardless of participation in SFA, were scoring on par with one another as a group. This phenomenon occurred with content-based tests (reading, mathematics, and language arts) and in the reading and listening and speaking portions of the California English Language Development Test, an English-language proficiency measure. The only statistically significant difference among student groups was that students in both ALA and SFA appeared to be scoring at a lower level on the California English Language Development Test writing portion than matched peers in the other three groups of interest (participants in ALA but not SFA, participants in both SEI and SFA, and participants in SEI but not SFA). Additional findings, theoretical and methodological issues, and implications for future research are featured.

Introduction

Evaluations of bilingual education programs that aspire to promote bilingualism and biliteracy are prevalent, yet respected evaluation research studies to guide instructional decisions for English language learners (ELLs) are surprisingly few. The need for systematic research on bilingual education

occurs at a crucial time for ELLs. The last two decades have marked considerable change toward federally based, accountability oriented educational policy, most recently the No Child Left Behind Act (2002). The goal of No Child Left Behind is to improve teacher effectiveness and to raise student achievement by mandating that “all children” attain high performance standards.

The No Child Left Behind Act (2002) requires knowledge and application of “scientifically based research” in the curricular areas of reading, mathematics, and science, and in instructional methods and strategies related to parent involvement, professional development, language instruction, and several other areas. As part of identifying what constitutes scientifically based research, the American Institutes for Research (2002) developed a useful three-pronged framework to help teachers determine if a program or practice is based on scientific research: (a) theoretical base of the program or practice; (b) evidence of effects; and (c) implementation and replicability.

The number of ELLs is growing nationally as well, adding pressure to identify suitable language support programs for this student population. In 2000–2001, about 4.6 million ELLs matriculated in the U.S. K–12 educational system, with about 79% reporting Spanish as their first language (Kindler, 2002). By the 2030s, language-minority students are expected to comprise 40% of the school-aged population in the country (Thomas & Collier, 2002). In California, the need and stakes are higher. One third of the nation’s ELLs are in California, numbering over 1.5 million, comprising one fourth of the public school K–12 population generally and one in three students in the elementary grades (Gandara & Rumberger, 2002). According to the 2000 census, about 1.1 million children aged 5–17 speak another language and do not speak English “very well” (U.S. Census Bureau, 2000). Of persons speaking “other” languages, Spanish is the most common spoken language (82%), followed by Vietnamese (3%), Chinese and Cantonese (2%), Hmong (2%), and several other languages (U.S. Department of Education, 2000–2001).

This paper provides a brief summary of research related to language acquisition programs for ELLs, theoretical and methodological issues in conducting research in this area, and findings from an ongoing study of a K–5 transitional bilingual education program, Academic Language Acquisition (ALA), for Spanish-speaking ELLs in a large urban school district in northern California. The study takes into consideration the presence of a well-established reading program, Success for All (SFA), when it is offered in addition to the bilingual education program. It also compares students in these programs with those who enrolled in Structured English Immersion (SEI). Findings and implications for future research are featured.

Research to Date

For decades, research on the effects of bilingual education programs has yielded often conflicting findings, serving as fodder to support or to contest varying points of view. Proponents of bilingual education recommend the utilization of the students' native language (L1) as a vehicle toward academic and linguistic development in English. English Language Development is formally introduced in kindergarten and is a foundational component at each grade level thereon. Opponents note that English is best learned through immersion in an English-language environment without supports in L1 or with minimal L1 instruction. While it is not possible to provide a comprehensive discussion of previous evaluation studies of bilingual education programs, what follows is a summary of relevant research findings to date.

Meta-analyses and summaries of research suggest that students who are taught initially in their L1 and then transition to English ultimately read as well or better in English than students taught only in English (August & Hakuta, 1997). ELLs enrolled in well implemented, supportive bilingual education environments experience higher gains and higher magnitude effect sizes on content-based tests in reading, language skills, mathematics, and total achievement when the tests are administered in English, in comparison with English learners in poorly implemented bilingual environments. Performance is even higher for ELLs tested in their respective L1s (Willig, 1985). Finally, students who receive at least some of their instruction in their L1 perform significantly better on standardized assessments than comparable students taught entirely in English (Greene, 1998).

Of more recent studies, there is evidence suggesting that participation in bilingual instruction may narrow the achievement gap for ELLs and their more English-proficient peers, although overall performance remains below the desired standard. In California, Parrish, Linqanti, Merickel, Quick, Laird, and Esra (2002) found that in the elementary grades, the greatest achievement for the Grade 2 cohort occurred in schools that offered bilingual instruction before Proposition 227 in 1998 or continued to offer bilingual instruction after Proposition 227. In addition, there was a slight narrowing of the achievement gap between native English speakers in comparison with ELLs and reclassified ELLs in those two types of schools, while in schools that never offered bilingual education, there was no reduction. The achievement gap, however, remains wide and increases at the higher grades.

Thomas and Collier (2002), in a 5-year study of students matriculating in five urban and rural school districts nationwide (1996–2001), found four programs—90–10 two-way bilingual immersion (or dual language), 90–10 transitional bilingual education, 50–50 one-way developmental bilingual education, and 50–50 two-way developmental bilingual education—are the only programs of eight major program types for English learners that help

students reach the 50th percentile on both nationally standardized tests in English Total Reading (e.g., Stanford Achievement Test [9th ed.] [SAT-9] [Harcourt Educational Measurement, 1997]) and in Spanish Total Reading (e.g., Spanish Assessment of Basic Education [2nd ed.] [SABE/2] [CTB/McGraw-Hill, 2002b]) in all subjects and that help maintain (or increase) that level of high achievement through the end of schooling. Thomas and Collier note that the fewest dropouts occur in these programs. They also note that to reduce the typical achievement gap between ELLs and native English speakers, bilingual education programs must be effective, well implemented, not segregated, and sustained for an extended period of time (5–6 years).

Theoretical and Methodological Considerations

Numerous methodological and theoretical problems plague research and evaluation studies involving ELLs. The variation in study findings emerges largely from small sample sizes, difficulties in identifying and preserving meaningful comparison groups, inappropriate methodology and/or statistical analyses, and the lack of a theoretical framework to guide the design and conduct of the study itself. Even meta-analytic studies and reviews of multiple studies (e.g., Dulay & Burt, 1978; Baker & de Kanter, 1981; Willig, 1985; Greene, 1998) yielded conflicting findings depending on the types of studies selected for consideration, the criteria for selection, and the criteria for secondary analysis of the studies themselves. These concerns yielded critical reviews of the methodological aspects of evaluations of bilingual education programs more generally (e.g., Meyer & Fienberg, 1992). Ultimately, advocates and opponents of bilingual education agreed on the following four criteria for designation of studies as “methodologically acceptable”: (a) Studies had to compare students in bilingual programs to a control group of similar students; (b) Design had to ensure that initial differences between treatment and control groups were controlled statistically or through random assignment; (c) Results were to be based on standardized test scores in English; and (d) Differences between the scores of treatment and control groups were to be determined by means of appropriate statistical tests.

Research on bilingual education can be used to inform policy (the “Research–Policy” paradigm). Cummins (1999) argued that if we disregard studies that do not meet the “methodologically acceptable” criteria, we risk losing the important contributions that can come from these studies. He proposes an alternative “Research–Theory–Policy” paradigm in which the connection between research and policy is mediated by theory. An explicitly defined theory is necessary to guide the development and implementation of a bilingual education program, as well as to guide subsequent study of that program.

Guided by the above literature, the evaluation team for the current study tried to design a study that met the four methodologically acceptable criteria and drew on the theoretical frameworks of the programs of interest. Specifically, the evaluation team identified ELLs whose initial level of English-language proficiency was comparable and examined how these students performed on standardized tests in English when instructed according to one of two possible models. Efforts were made to control for initial differences through analyses of matched samples. Although these efforts to improve the theoretical and methodological quality of the study were not entirely successful, as discussed later in the Limitations of the Present Study section, the study was likely more rigorous than it otherwise would have been.

Purpose of Study

This study examined the effectiveness of the transitional bilingual ALA program in enhancing K–5 students’ English-language proficiency, as well as their English performance in academic subject areas, in comparison with the SEI process. An existing reading program in some schools, SFA, served as a confounding influence because it had similar goals for reading development and included English- and Spanish-language curricula. Several questions were addressed:

1. What are the criteria to determine the effectiveness of the ALA program and SEI process?
2. Are ALA and SEI students reaching their stated programmatic goals?
3. How do the ALA program and the SEI process compare in helping students increase their English-language proficiency, in general and over time?
4. What impact does the presence of an SFA program have on students’ English-language proficiency, when combined with ALA or SEI?

Program/Process Descriptions

The following descriptions of the ALA program and SEI process come from the July 1999 “Master Plan for English Learners” prepared by the school district considered in the present study.

Academic Language Acquisition

This transitional bilingual education program offers English-language development and content instruction at grade level for Spanish-speaking ELLs. Instruction is in Spanish and English, with increasing English-language instruction over time, along with access to the core curriculum. In kindergarten and first grade, 70% of instruction is provided in the students’ L1 and 30% in

English (L2). As students progress through each grade, the percentage of English increases until the fourth and fifth grades, when the program design calls for 85% English-language and 15% Spanish-language instruction.

Specific conditions must be met to trigger an ALA program implementation: (a) There are 20 or more Spanish-speaking ELLs in a school's attendance area for each grade K–5; (b) Those students have, pursuant to informed consent, requested a parental waiver and chosen to participate in the ALA program; and (c) There are sufficient certified Spanish bilingual teachers to instruct the students. As a result, ALA tends to be implemented in neighborhoods where a critical mass of Spanish-speaking ELLs resides to maximize the impact of a fully articulated, sequential bilingual program.

The ALA bilingual program assists ELLs with “less than reasonable fluency” in developing academic skills in the core curriculum using L1 and L2 instruction. Less than reasonable fluency is defined as scoring at Levels 1–3 on the Language Assessment Scales (LAS) or below the “early advanced” level on the California English Language Development Test (CELDT) (CTB/McGraw-Hill, 2002a). To maximize students' development of cognitive academic-language proficiency in L1 and L2, teachers provide instruction through scheduled language blocks throughout the school day. The goals of the ALA program are English-language proficiency and academic achievement in the core curriculum utilizing L1. There are three types of achievement goals: academic, linguistic, and psychosocial, although no psychosocial goals were included in the school district's expectations or descriptions. The district's goals for ALA were:

1. Students will read, write, and perform mathematics processes on grade level by third grade in their L1 (academic goal).
2. Students will be orally fluent in English by third grade (linguistic goal).
3. Students will transition into English reading by the end of third grade (academic goal).

Structured English Immersion

Spanish-speaking ELLs whose parents do not request the ALA program or for whom there is not an established bilingual program in the school instead enroll in a mainstream English-language classroom, a phenomenon known as SEI. In SEI, ELL students receive specialized instruction, such as English Language Development and Specially Designed Academic Instruction in English or sheltered English strategies, geared to their individual educational and linguistic needs and performance. Students receive access to the core curriculum, as they become increasingly fluent in the English language (the literature does not specify a target English fluency level). SEI is implemented in both ALA and non-ALA schools. Further, SEI classes enroll both students with “less than reasonable fluency” (LAS 1–3 or CELDT intermediate and below), and students with “reasonable fluency” (LAS 4–5 or CELDT early

advanced and higher). Within the same school, however, the number of students in SEI classes tends to be disproportionately lower than students enrolled in ALA classes. In schools where ALA is offered, it is possible for some English learners to receive SEI instruction if there are not enough students to comprise an ALA classroom (a minimum of 20 students is necessary).

In SEI, L1 support is provided to facilitate English acquisition and academic progress. Support may occur in a variety of ways, such as through bilingual staff, tutors (e.g., parent and community volunteers, school district paid tutors, college students, or peer tutors), paraprofessionals, and instructional associates. Known as the “sequential method,” this process provides appropriate language assistance programs that are consistent with state and federal guidelines for elementary Spanish-speaking ELL students not enrolled in the ALA program. Daily instruction is predominately in English, with about 10% to 15% of instruction in English Language Development, 75% using sheltered or Specially Designed Academic Instruction in English strategies. The remaining 10% to 15% is instruction in L1, Spanish. This instructional allocation remains generally consistent across the K–5 grade levels.

Although ALA and SEI have the same long-term goal of improving English-language acquisition among ELLs, their intermediate goals differ based on the different instructional philosophies. One goal, however, is the same for both ALA and SEI: Students will be orally fluent in English by third grade (linguistic goal). The following are all the goals for SEI, as denoted by the school district:

1. Students will read, write, and perform mathematical processes in English on grade level by third grade (academic goal).
2. Students will be orally fluent in English by third grade (linguistic goal).
3. Students will meet the redesignation criteria from limited English proficient status to fluent English proficient status at the following rate: 25% of students at the end of third grade, 50% at the end of fourth grade, 100% at the end of fifth grade (academic goal).

Success for All

SFA is a nationally known, widely implemented, schoolwide, coordinated K–8 reading program² for elementary schools. Offered in eight schools with the ALA program and/or SEI process, particularly those with lower socioeconomic status levels, SFA represents a major confounding influence when examining program effects. Founded on research-based cooperative learning strategies, SFA groups students according to language and reading level. Students engage in 90 minutes of uninterrupted daily reading instruction and read specially designed curricula customized by reading level. (For an overview of SFA and a summary of research, see Slavin & Madden, 2001.)

SFA is based on several principles: (a) emphasis on prevention, early and intensive intervention, and tutoring for students with academic difficulties; (b) emphasis on the integration of phonics and meaning-focused instruction, cooperative learning, and curriculum-based assessments; (c) kindergarten instruction with storytelling and language development; (d) adaptations for Spanish and English as a Second Language; (e) a family support program engaging parents, community members, and integrated service; and (f) extensive professional development for teachers through the elementary grades.

At least 80% of the teaching staff within a school must approve SFA implementation before the program can be adopted at that school. Although the adopted components vary by site, depending on school needs and available resources, several core scaffolding elements are common to all implementations (Slavin & Madden, 2001). The common elements include: (a) a reading program, whereby students engage in a regular 90-minute reading period, grouped based on their reading performance levels into reading classes of students at the same level; (b) 8-week reading assessments designed by SFA administered to document student progress and, in conjunction with teacher judgment, to determine which students might require additional tutoring or other supports, to rearrange students' reading groups, and to identify additional adaptations necessary to suit students' needs; (c) reading tutors for one-on-one instruction with students who might be having difficulties in their reading groups (tutoring occurs outside of the students' regular reading or mathematics periods for 20-minute sessions); (d) a family support team, consisting of the Title 1 parent liaison, vice principal (if any), counselor (if any), and other appropriate staff, which works in a variety of ways to strengthen the home-school connection; (e) a program facilitator, who oversees the overall implementation at his or her school; and (f) teachers and teacher training, whereby SFA teachers are certified and receive detailed teacher's manuals, in-service, and professional development to address instructional pace, cooperative learning strategies, and overall effective implementation of designated SFA curricula.

Theory Guiding Research

The two programs of interest in this study, ALA and SEI, are founded on theoretically grounded research (see below), and to a great extent they are based on the same pedagogical models. The key difference is the extent to which the English language is used in the instructional process. We were not able to address all features of ALA and SEI in the present study, particularly parental involvement, teacher quality, and increasing minority status.

Length of Time in Program

Previous research suggests that it takes 4–6 years for ELLs to achieve success in English-language acquisition generally and within core content areas (Cummins, 1979, 1981). We hypothesized that students' English-language proficiency would improve over time, in conjunction with greater exposure to and experience with the English language.

Adequate Exposure to Second Language

In order to acquire an L2, children need sufficient exposure to that language in both formal (teacher-directed) and informal (with friends) situations. The type of exposure is as important as the amount (Krashen, 1982, 1985). In this study, both programs offer L2 support in English; the difference lies in how much instructional exposure these students receive in English, and the quality of the instruction itself.

Emphasis on Academic Achievement

In addition to English-language development, the evaluation team in this study was interested in students' acquisition of core content knowledge in academic subjects (e.g., mathematics) (Cummins, 1981). We would expect that a successfully implemented language-acquisition program would promote and achieve not only proficiency in English but also knowledge of academic content in the English language.

Bilingual Instruction Through Separate Monolingual Lesson Periods

Separation of languages for instruction helps to promote language development (Baker & de Kanter, 1981; Dulay & Burt, 1978; Swain, 1983).

Program Quality

Exposure to the English language, by itself, is not sufficient to encourage language development. Positive and reciprocal interactions between students and teachers often result in improved educational outcomes for students (Kagan, 1986). We would expect that programs that offer high-quality language exposure, academically and socially, along with multiple and creative opportunities to actively engage students in the learning process, are more effective at facilitating language acquisition than programs that do not.

Type of English-Language Instruction

This is the primary instructional feature that differentiates the ALA and SEI approaches. The ALA program promotes bilingual instruction through separate lesson periods in the students' L1, Spanish. As students progress in their schooling, their exposure to L1 decreases, while instruction in English increases. In contrast, the SEI model promotes earlier and greater exposure to

the English language (in part related to the “time-on-task” hypothesis, which states that more time spent on English may lead to increased fluency in English), using Specifically Designed Academic Instruction in English and sheltered teaching strategies, with less overall exposure to the students’ L1.

Instructional Environment

The link between the student and teacher more broadly is an important influence on students’ learning and growth. Students who are comfortable, confident, efficient, and motivated are more active, positive learners than students who lack these qualities. Classroom and school environments that foster these qualities are more conducive to student learning than schools that do not.

Parental Involvement

Parents who actively participate in educational activities, who take an active interest and role in their child’s education, and who support the teacher’s instructional approach are more likely to positively influence their children than parents who do not (Cummins, 1986).

Teacher Quality

We would expect that teachers familiar with key instructional strategies, such as cooperative learning, L2 acquisition, and grade-level standards and curricula, and who embody excellent teaching skills and knowledge, would be better equipped in developing positive, encouraging relationships with their students and their learning processes. For ELLs, teachers with native-language fluency in Spanish, their students’ L1, can often communicate more meaningfully. Some notions of quality may be captured based on the teachers’ training credentials and teaching experience.

Increasing Minority Status

There is general agreement that being bilingual and bicultural are positive attributes in an increasingly culturally diverse country. Programs that can effectively promote all students’ self-esteem and cultural appreciation can often enhance student learning (Cummins, 1986).

Methodology

Participants and Procedures

The study focused on 829 K–3 Spanish-speaking ELL students in a large urban school district in northern California. Participants were enrolled as kindergartners in 1998–1999 in ALA or SEI classes and remained in that particular

type of class through the first grade in 1999–2000, second grade in 2000–2001, and third grade in 2001–2002; they had begun the fourth grade in 2002–2003, when the study was conducted. Students transferring between schools remained in the study sample as long as they continued in the same type of class in which they started.

An evaluation team, which consisted of two university professors and two graduate students, worked closely with school district officials to develop criteria for evaluating implemented ALA and SEI ventures. Criteria for evaluating implementation for both ALA and SEI were similar, focusing on the following factors: (a) staff understanding of foundations of bilingual education or English immersion, depending on the program, (b) staff articulation of foundations as related to the program, (c) staff experience related to implementing the program, (d) whether the program was fully articulated across grade levels in school, (e) administrative and teacher leadership on site, (f) whether the staff identified itself as part of the program, (g) whether parental support was an important part of the program, and (h) whether compliance criteria were met.³ Ultimately, all ALA and SEI schools in the district were considered to have met these criteria and were included in the study.⁴

The evaluation consisted of several portions. The evaluation team visited the ALA program and SEI process schools, conducting interviews with key stakeholders (e.g., school principal, instructional staff, teachers); classroom observations at Grades 1 and 3, focusing on program implementation factors (e.g., staff understanding of bilingual education); and analysis of program documents available from the school district, including teacher training materials. Student demographic and test score data were obtained, such as the type of class in which ELLs were enrolled; and school enrollment and attendance data. Report drafts were circulated among stakeholders to ensure accuracy.

Instruments

Several assessment measures were used as outcome measures.

California English Language Development Test

In 2001–2002, the CELDT replaced the LAS oral tests, the primary indicator of English proficiency to identify ELLs. The CELDT provides weighted scores for the skill areas of listening, speaking, reading, and writing, as well as an overall score. This test is aligned with the California State English Language Development content standards.

Stanford Achievement Test (9th ed.), Form T

The SAT–9 is a nationally normed test administered annually to students in Grades 2–11, including reading, mathematics, and language arts (written expression).

Spanish Assessment of Basic Education (2nd ed.)

Normed with a national sample of Spanish-speaking students, this test is administered to all Spanish-speaking ELLs who are “newcomers” to the United States (essentially, students enrolled in California public schools for less than 12 months prior to testing). The SABE/2 is administered annually to students in ALA, but not SEI, classrooms.

Findings

Evaluations of bilingual education programs are often criticized due to the lack of definable comparison groups, thus making it difficult to compare students within and across different types of programs or because of inappropriate study designs and analyses (see, for example, Dannoff, Coles, McLaughlin, & Reynolds, 1978; Development Associates, 1986; Ramírez, Pasta, Yuen, Billings, & Ramey, 1991). In analyzing the current data, the evaluation team tried to minimize such problems. Rather than include all students, we chose to analyze data only from students who met designated criteria: (a) enrolled continuously in an ALA program or in an SEI classroom, with or without SFA, since kindergarten (1998–2002);⁵ (b) specified ethnicity as “Hispanic” and/or “home language” as Spanish; (c) high attendance, meaning present for up to 182 days⁶ with a maximum of 15 days absent (at least 167 days present); and (d) longitudinal test data available. The final sample consisted of 441 students, of whom 363 (82%) were enrolled in ALA classes and 78 (18%) were in SEI classes in their respective schools.⁷ Taking the presence of the SFA programs into account, the percentage breakdowns for the following data analyses were: students in both ALA and SFA, 39%; students in ALA but not SFA, 43%; students in both SEI and SFA, 6%; and students in SEI but not SFA, 12%.

Content-Area Assessment in Primary Language (Academic Language Acquisition Students Only)

ALA students scored in the upper percentiles on the SABE/2 tests. In third grade, students were receiving 65% of their classroom instruction in Spanish and tended to perform better on tests that matched their language of instruction. ALA students with SFA scored similarly in reading and mathematics, but lower in language arts, than did ALA students without SFA instruction. In reading, the median national percentile ranks (NPRs) for ALA students with and without SFA were 67 and 68, respectively. ALA students’ performance in mathematics was the highest, with a median NPR of 78 regardless of SFA participation. The biggest difference between SFA and non-SFA students was in language arts, where students in ALA but not SFA (median NPR = 78) scored higher than students in both ALA and SFA (median NPR = 67).

One can also examine student performance based on the percentage of students performing at or above a designated quartile (25th, 50th, or 75th percentiles) or based on the percentage scoring at or above the national median. In general, ALA students scored well on the SABE/2 relative to the national norming sample. In reading, 87% of ALA students with SFA and 92% of ALA students without SFA scored at or above the national median. Similar trends occurred in mathematics (with students in both ALA and SFA scoring 86%, and students in ALA but not SFA scoring 91%) and language arts (students in both ALA and SFA scored 81%, and students in ALA but not SFA scored 90%).

Measures of English-Language Proficiency

The CELDT examines listening and speaking, reading, and writing, categorizing students into five levels in terms of skill-area proficiency: beginning, early intermediate, intermediate, early advanced, and advanced. The percentages of students scoring above and below the early advanced proficiency level were compared, because this level is an indicator for possible reclassification into “fluent English proficient” status (see Table 1). Students performed best on the listening and speaking portion, with 27% to 38% of ALA students and 58% to 67% of SEI students scoring at or above the early advanced level, with and without SFA, respectively. About 9% of students in both ALA and SFA, 17% of those in ALA but not SFA, 32% of those in both SEI and SFA, and 41% of those in SEI but not SFA scored at the early advanced level or above in writing. Performance of all groups was somewhat lower in reading.

Content-area assessment in English

In general, students performed better on the SAT–9 mathematics tests than on the reading or language arts tests, most likely due to lower English-language proficiency demands (see Table 2). The median NPRs in mathematics ranged from 39 (students in both ALA and SFA) to 47 (students in ALA but not SFA) for ALA students and 42 (SEI but not SFA) to 57 (both SEI and SFA) for SEI students. Students in ALA scored lower on the SAT–9 in reading than those in SEI. This finding is expected, however, since ALA students received only 40% of their instruction in English (60% in Spanish) at this time, compared with 90% English instruction for SEI students. As the SAT–9 is an English-language test, it is not surprising, therefore, that ALA students scored lower than SEI students. In contrast, ALA students scored higher on the Spanish-language SABE/2 than on the English-language SAT–9.

Interestingly, ALA students who participated in the SFA program scored slightly higher on the SAT–9 than ALA students not in SFA. The opposite occurred with SEI students. In reading, about 20% of students in both SEI and SFA, and 31% of those in SEI but not SFA, scored at or above the national

Table 1

Percentage of Students Scoring at Each Level on California English Language Development Test

	ALA and SFA (n = 171)	ALA but not SFA (n = 147)	SEI and SFA (n = 25)	SEI but not SFA (n = 39)
Listening & speaking portion				
Below early advanced	73	62	42	33
Early advanced & above	27	38	58	67
Reading portion				
Below early advanced	88	89	95	90
Early advanced & above	12	11	5	10
Writing portion				
Below early advanced	91	83	68	59
Early advanced & above	9	17	32	41
Total				
Below early advanced	87	74	58	51
Early advanced & above	13	26	42	49

Note. ALA = Academic Language Acquisition; SEI = Structured English Immersion; SFA = Success for All. Analysis does not control for initial differences among groups. Sample consists of beginning fourth-grade students who enrolled in ALA or SEI for 4 consecutive years (1998–2002), who demonstrated good attendance (at least 167 days present per academic year), and for whom test data were available.

Table 2

Percentage of Students at or Above Stanford Achievement Test (9th ed.) National Percentile Ranks (NPRs)

	ALA and SFA (n = 170)	ALA but not SFA (n = 189)	SEI and SFA (n = 25)	SEI but not SFA (n = 52)
Reading portion				
75th NPR	2	0	8	6
50th NPR	14	9	20	31
25th NPR	41	44	68	62
Mathematics portion				
75th NPR	19	20	32	27
50th NPR	38	46	40	56
25th NPR	69	80	84	83
Language arts portion				
75th NPR	4	5	8	14
50th NPR	19	20	40	40
25th NPR	46	52	88	75

Note. ALA = Academic Language Acquisition; SEI = Structured English Immersion; SFA = Success for All. Analysis does not control for initial differences among groups. Sample consists of third-grade students enrolled in ALA or SEI for 4 consecutive years (1998–2002), who demonstrated good attendance (at least 167 days present per academic year), and for whom test data were available.

median for that year, compared with students in both ALA and SFA (14%) and those in ALA but not SFA (9%). In mathematics, ALA students performed comparably to SEI students, particularly students in ALA but not SFA. ALA and SEI students without SFA performed highest on the SAT–9 mathematics test (46% and 56%, respectively), compared with 38% for those in both ALA and SFA and 40% for those in both SEI and SFA. Finally, students' performance on the SAT–9 language arts test was better than in reading. SEI students outperformed ALA students. About 19% of students in both ALA and SFA

and 20% of those in ALA but not SFA scored at or above the 50th national percentile. In comparison, 40% of SEI students, with or without SFA, scored at this level.

Matched Samples

Efforts were made to obtain a more fair comparison of the four groups of interest. The smallest comparison group (students in both SEI and SFA) included 25 students and represented the base group. Twenty-five students in each of the other groups (both ALA and SFA, ALA but not SFA, and SEI but not SFA) were selected, based on initial English-language proficiency in kindergarten and first grade using Pre-LAS assessments. In all, the matched sample reflected a total of 100 students, although not all students had complete test data (which explains why some subgroup sample sizes may be less than 25). Results were compared for similarity and recalibrated.

Measures of English-language proficiency

After 4 years in ALA or SEI, students' CELDT performance in English-language proficiency was examined (see Table 3). About half of the students in the four groups of interest scored at or above the early advanced level in listening and speaking. Scores in reading were somewhat lower. Significant differences among groups were not found in listening and speaking ($F[3,69] = .758, p = .522$) or reading ($F[3,69] = .787, p = .505$), but rather in writing ($F[3,69] = 4.18, p = .009$). Students in both ALA and SFA scored at a lower CELDT writing level ($M = 2.21$) than students in the other three groups (ALA but not SFA, $M = 2.89$; both SEI and SFA, $M = 3.00$; SEI but not SFA, $M = 3.13$).

Content-area assessment in English

The matched sample comparison yielded trends similar to those seen in the aggregated student data (see Table 4). Analyses of variance, including computation of pairwise contrasts, were conducted to determine if there were significant differences in mean SAT-9 scores between matched student groups as of the end of the third grade. Based on these analyses, no statistically significant differences were found in reading ($F[3,96] = .45, p = .7126$), in mathematics ($F[3,96] = .23, p = .87$), or in language arts ($F[3,96] = 1.12, p = .34$).

Given limitations of the matched data, one issue is whether the sample sizes ($n = 25$ in each group) were large enough to expect that statistically different means would be found if the groups were in fact different. To assess this, effect sizes given standardized units of measurement were computed for a priori pairwise comparisons. Of the 57 computations, 56 had effect sizes below .20, or "small" effects (Cohen, 1988). Another issue is multiple testing. At $p = .05$, we would expect to find 5 of 100 associations statistically significant when in fact there are no associations among the variables. In this instance, 5 of 57 associations were significant, slightly more than we would expect to find by chance (Cohen).

Table 3

Matched Sample—Percentage of Students Scoring at Each Level on California English Language Development Test

	ALA and SFA (n = 19)	ALA but not SFA (n = 19)	SEI and SFA (n = 19)	SEI but not SFA (n = 16)
Listening & speaking portion				
Below early advanced	42	32	42	44
Early advanced & above	58	68	58	56
Reading portion				
Below early advanced	84	89	95	94
Early advanced & above	16	11	5	6
Writing portion				
Below early advanced	95	74	68	75
Early advanced & above	5	26	32	25
Total				
Below early advanced	68	37	58	63
Early advanced & above	32	63	42	37

Note. ALA = Academic Language Acquisition; SEI = Structured English Immersion; SFA = Success for All. Matched sample consists of beginning fourth-grade students who enrolled in ALA or SEI for 4 consecutive years (1998–2002), who demonstrated good attendance (at least 167 days present per academic year), and for whom test data were available.

Table 4

Matched Sample—Percentage of Students at or Above Stanford Achievement Test (9th ed.) National Percentile Ranks (NPRs)

	ALA and SFA (n = 25)	ALA but not SFA (n = 25)	SEI and SFA (n = 25)	SEI but not SFA (n = 25)
Reading portion				
75th NPR	8	0	8	0
50th NPR	24	12	20	28
25th NPR	56	64	68	60
Mathematics portion				
75th NPR	25	24	32	32
50th NPR	49	60	40	52
25th NPR	81	88	84	80
Language arts portion				
75th NPR	12	0	8	12
50th NPR	36	20	40	40
25th NPR	48	64	88	76

Note. ALA = Academic Language Acquisition; SEI = Structured English Immersion; SFA = Success for All. Matched sample consists of beginning fourth-grade students who enrolled in ALA or SEI for 4 consecutive years (1998–2002), who demonstrated good attendance (at least 167 days present per academic year), and for whom test data were available.

Conclusion

Attempts were made to compare students in a transitional bilingual education program (ALA) with comparable students in a control group (SEI). Several outcome variables of interest were used, most notably standardized test scores from standards-based assessments administered in English. Efforts to control for initial differences among the groups were made through a matching procedure. Data were analyzed through descriptive procedures, analysis of variance, and calculation of effect sizes to determine if differential effects were statistically significant.

After 4 years in their respective programs, students in ALA and SEI classes displayed only nominal differences, at best, in their performance on various achievement indicators. ALA and SEI students, regardless of participation in SFA, were comparable on English-language SAT–9 tests in reading, mathematics, and language arts, as well as the reading and listening and speaking portions of the CELDT, an English-proficiency test. The only significant difference among groups occurred in writing, where students in both ALA and SFA scored lower than their peers. It is premature to make inferences about the role of the SFA program on students' performance, however, particularly for ALA students. One could hypothesize that students' exposure to two well-established programs—one focusing on English-language acquisition, the other focused on improving reading skills—could have potentially negative effects, although additional investigation is needed.

Limitations of the Present Study

Many theoretical and methodological problems encountered in earlier evaluations of bilingual programs were also encountered in this study, despite efforts to overcome them. In the following section, such limitations are discussed, along with their implications for the current study.

Small sample size

One of the problems in multi-site evaluation studies is treatment fidelity or ensuring that there are well-defined treatment and control groups. The current study imposed multiple criteria to better define the groups of interest for comparison purposes. These criteria related to students' attendance rates, matriculating in the same program, and being enrolled for 4 consecutive years reduced the final numbers to 441 total students in the sample ($n = 363$ ALA; $n = 78$ SEI). Less than one fifth of the students (18%) were SEI participants, and they were scattered across several schools.

Student mobility

Students who stayed within ALA or SEI but who changed schools were not differentiated from students who stayed in the same program at the same school. Thus, even though the same program may have been offered, the quality of the program implementation varied, as well as teachers, other instructional modalities, and overall classroom and school environments.

Lack of multiple measures on student achievement

Questions about the validity of inferences on standardized tests, administered in the English language, for ELLs are well documented (see, e.g., August & Hakuta, 1997). Even though multiple sources of evidence are necessary to make suitable inferences about students, only standardized test scores were available for this study. Future studies should include additional indicators of academic achievement, as well as more process-oriented, qualitative data that examine the context in which the test scores were acquired.

Lack of information about program implementation

It is important to ensure that the program is implemented as intended. Often, due to circumstances beyond the control of the evaluators or of the school district staff (involving, e.g., school resources, management, access to information, program support), implementation of a program may be less than ideal.

Difficulty in making causal inferences about program effects

Because of potentially confounding variables (e.g., after-school educational programs, parental tutoring), potential measurement problems, lack of information about program implementation across schools, and other caveats, it is virtually impossible to make causal inferences about effects of the ALA program and/or SEI process on student performance. The best that we can do is to accumulate evidence of a connection between observed phenomena; ultimately, more evidence across circumstances and student populations will give greater credence to a strong inferential connection.

Implications of the Present Study

Given findings to date, several theoretical contributions of this study may be useful for informing instructional decisions for ELLs. First, ELLs in a transitional bilingual education program need several years (at least 4 full years) to reach the English-language attainment levels of their peers in English-immersion classes. Simultaneously, students in the ALA program appear to be maintaining their L1 proficiency on the content-based tests of interest. Second, the emphasis on academic achievement of both ALA and SEI instruction appeared to be yielding positive effects, although inconclusive findings with students enrolled in SFA suggest that participation in two academically oriented programs may yield conflicting outcomes and thus potentially decreased student performance. Third, the level of program implementation (which encompasses teacher quality, emphasis on academic achievement, quality student and teacher interactions, active parental involvement, etc.) is an important consideration that we were not able to examine as closely as desired. However, given that potentially “developing” ALA and SEI ventures were included, it is likely that their presence resulted in an underestimation of student performance. Finally, this study may be used to test the “time-on-task” hypothesis, that is, a primary focus on English instruction and English educational curriculum, particularly in the early years, will address the educational achievement needs of ELL children. The results of the present study, however, do not support this hypothesis, as it appears that after 4 years in the ALA program, ELLs performed comparably to their SEI peers on selected English-language tests. While the current study indeed encountered numerous limitations, it does feature key findings that provide important evidence about how Spanish-speaking ELLs may perform when exposed to one of two different instructional modalities.

References

- American Institutes for Research. (2002). *Evidence of effects on student achievement*. Unpublished manuscript, U.S. Department of Education, Washington, DC.
- August, D., & Hakuta, K. (1997). *Improving schooling for language-minority children: A research agenda*. Washington, DC: National Academy Press.
- Baker, K. A., & de Kanter, A. A. (1981). *Effectiveness of bilingual education: A review of the literature*. Washington, DC: Office of Planning, Budget, and Evaluation, U.S. Department of Education.
- Cohen, J. (1988). *Statistical power analysis for the behavioral science* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- CTB/McGraw-Hill (2002a). *California English Language Development Test*. Monterey, CA: Author.
- CTB/McGraw-Hill (2002b). *Spanish Assessment of Basic Education* (2nd ed.). Monterey, CA: Author.
- Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research*, 19, 222–251.
- Cummins, J. (1981). The role of primary language development in promoting educational success for language minority students. In California State Department of Education (Ed.), *Schooling and language minority students: A theoretical framework* (pp. 3–49). Los Angeles: National Dissemination and Assessment Center.
- Cummins, J. (1986). Empowering minority students: A framework for intervention. *Harvard Educational Review*, 56(1), 18–35.
- Cummins, J. (1999). Alternative paradigms in bilingual education research: Does theory have a place? *Educational Researcher*, 28(7), 26–32.
- Dannoff, M. N., Coles, G. J., McLaughlin, D. H., & Reynolds, D. J. (1978). *Evaluation of the impact of ESEA Title VII Spanish/English bilingual education programs* (Vol. 3). Palo Alto, CA: American Institutes for Research.
- Development Associates. (1986). *Year 1 report of the longitudinal phase*. Arlington, VA: Author.
- Dulay, H. C., & Burt, M. K. (1978). *Why bilingual education? A summary of research findings* (2nd ed.). San Francisco: Bloomsbury West.
- Gandara, P., & Rumberger, R. (2002, December). *What all of us can and must do to address the achievement gap for English learners*. Paper presented at the Accountability Institute: English Learners and Immigrant Students, Santa Barbara, CA.

- Greene, J. (1998). *A meta-analysis of the effectiveness of bilingual education*. Claremont, CA: Tomas Rivera Policy Institute.
- Harcourt Educational Measurement. (1997). *Stanford Achievement Test* (9th ed.), Form T. San Antonio, TX: Author.
- Kagan, S. (1986). Cooperative learning and sociocultural factors in schooling. In *Beyond language: Social and cultural factors in schooling language minority students*. Los Angeles: Evaluation, Dissemination, and Assessment Center, California State University.
- Kindler, A. L. (2002). *Survey of the states' limited English proficient students and available educational programs and services, 2000–2001 summary report*. Washington, DC: National Clearinghouse for English Language Acquisition and Language Instruction Educational Programs.
- Krashen, S. (1982). *Principles and practice in second language acquisition*. New York: Pergamon Press.
- Krashen, S. (1985). *Input in second language acquisition*. Oxford, England: Pergamon Press.
- Meyer, M. M., & Fienberg, S. E. (Eds.). (1992). *Assessing evaluation studies: The case of bilingual education strategies*. Washington, DC: National Academy Press.
- No Child Left Behind Act of 2001, Pub. L. No. 107-110, 115 Stat. 1425 (2002).
- Parrish, T., Linqianti, R., Merickel, A., Quick, H., Laird, J., & Esra, P. (2002). *Effects of the implementation of Proposition 227 on the education of English learners, K–12*. Palo Alto, CA: American Institutes for Research.
- Ramírez, D. J., Pasta, D. J., Yuen, S. D., Billings, D. K., & Ramey, D. R. (1991). *Final report: Longitudinal study of structured-English immersion strategy, early-exit and late-exit transitional bilingual education programs for language minority children* (Vol. 2). San Mateo, CA: Aguirre International.
- Slavin, R. E., & Madden, N. A. (Eds.). (2001). *Success for All: Research and reform in elementary education*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Swain, M. (1983). *Communicative competence: Some roles of comprehensible input and comprehensive output, in its development*. Paper presented at the 10th University of Michigan Conference on Applied Linguistics, Ann Arbor, MI.
- U.S. Census Bureau. (2000). *Summary File 3, Table P19. Age by language spoken at home by ability to speak English for the population 5 years and over*. Washington, DC: U.S. Government Printing Office.

- U.S. Department of Education. (2000–2001). *Survey of the states' limited English proficient students and available educational programs and services*. Washington, DC: U.S. Government Printing Office.
- Thomas, W. P., & Collier, V. P. (2002). *A national study of school effectiveness for language minority students' long-term academic achievement*. Santa Cruz: Center for Research on Education, Diversity, & Excellence, University of California, Santa Cruz.
- Willig, A. C. (1985, Fall). A meta-analysis of selected studies on the effectiveness of bilingual education. *Review of Educational Research*, 55(3), 269–317.

Endnotes

- ¹ Unlike ALA, a developed *program*, SEI is an instructional *process*. It does not have a well-articulated developmental framework like the ALA program.
- ² The SFA program actually serves preK through eighth-grade students. In the school district in the present study, however, it is implemented K–8 only.
- ³ This study was conducted as part of a broader federal desegregation court case. Compliance criteria for ALA programs were reviewed in a validation study conducted by the federal court compliance monitor. These criteria included teacher certification requirements and program waiver requirements.
- ⁴ Short-term case studies of three schools were also prepared to provide a more in-depth examination of the respective programs, although these are not presented in this paper.
- ⁵ If school data were missing for a student for a given year, it was assumed that the student was not enrolled for that year. The student was then excluded from the sample.
- ⁶ Attendance data were available for only academic year 1999–2000.
- ⁷ SEI classes include ALA-eligible students in ALA schools, but in SEI classrooms.

