

Reading Trajectories of Immigrant Latino Students in Transitional Bilingual Programs

Leslie Reese

California State University, Long Beach

Ronald Gallimore and Donald Guthrie

University of California, Los Angeles

Abstract

Using data from a random sample of Latino students in California, tracked throughout their elementary and middle school years, this paper examines ways in which outcomes vary for students of similar language and socioeconomic background who are initially instructed in their native language in transitional bilingual programs. As the students made the transition to instruction in English, which took place for most students in the sample between Grades 2 and 4, all students experienced an abrupt decline in performance on standardized reading test scores in English. However, performance trajectories after transition took markedly different paths, with higher achievers returning to pre-transition rates of performance as lower achievers continued an achievement decline that began for many while still being instructed in Spanish. The paper also examines school factors potentially associated with variations in performance (school-wide reform and strong kindergarten program) for subsets of the sample.

Introduction

There are reports of a “temporary lag” in achievement when students in Canadian bilingual programs are instructed in the second language (Baker, 1996). However, in Canadian-type immersion programs and dual-immersion programs in the United States, this temporary delay is reported to be overcome by sustained high levels of performance in later years (Baker, 1996; Thomas & Collier, 1997, 2002). Ramírez (1992) found similar longer term benefits associated with primary-language instruction in the late-exit bilingual programs under study. The recovery is a predictable outcome of the bilingual instructional

theory that specifies an interdependence of literacy-related and academic skills across languages (Cummins, 1991). The moderate to strong correlation between academic skills in the primary language and in the second language are manifestations of a common underlying proficiency enhanced by continued primary-language instruction (Cummins, 2001). Thus, the general assumption is that although transition to second-language instruction may produce a lag in performance, recovery will follow due to the hypothesized underlying proficiency.

However, in terms of individual English language learners, there is limited evidence regarding the lag and recovery. One reason is that many studies of bilingual education have focused on comparing outcomes across programs. Another reason is a lack of studies of individuals' performance trajectories over time. With few exceptions (Reese, Garnier, Gallimore, & Goldenberg, 2000; Siantz, 1997; Suárez-Orozco & Suárez-Orozco, 1997), studies of Spanish-speaking children's achievement are most often cross-sectional, correlational designs. As a result, little is known about the nature and magnitude of the lag in reading performance, the variance within English language learner populations, and the factors associated with the lag and with the predicted recovery.

This report was designed to address some of the empirical holes in our knowledge of the performance lag. It takes advantage of a longitudinal study of Latino youth, the UCLA Home-School Project (Goldenberg, Gallimore, Reese, & Garnier, 2001; Reese et al., 2000). The design of this longitudinal project, which included collection of standardized test scores and teacher ratings of student performance on a yearly basis and maintained a 75% participant retention rate, permitted us to examine the impact on reading performance of transition from bilingual to English-only instruction.

It is important to emphasize that the research design does not permit an evaluation of the bilingual program that the study participants received. Since all participants were initially placed in transitional bilingual programs, there was no basis of comparison. Second, the longitudinal project did not collect data on program quality; rather, it sampled from two school districts that were deploying transitional bilingual programs. Thus, the unit of analysis in the study reported here is individual student trajectory over time, before and after transition from Spanish to English instruction.

The longitudinal study has already published results on the ways in which family practices, some of which are rooted in home country experiences, support literacy development in the study sample and serve to explain part of the variation in outcomes for the second-generation children. However, children's year-to-year achievement trajectories remained unexamined. Of additional interest is the fact that the study had been designed in such a way that students' performance at schools in which specific interventions were under way—a well-implemented kindergarten literacy program at one school

site and a school-wide language arts reform effort at another site—could be compared with the performance of students at two different sites in two different districts without such interventions. Thus, the focus in the present study is on the ways in which transition from Spanish to English literacy instruction impacts student achievement, variations within the sample, factors associated with post-transition lag and recovery, and the effects of different school interventions on performance.

Research Questions

1. How well do Spanish-speaking children, initially placed in transitional bilingual instructional programs, perform in Spanish and English reading in elementary through middle school?
2. Overall, do students exhibit similar trajectories of recovery in English-reading proficiency after transition to English, regardless of initial literacy performance?
3. Are the instructional interventions (a school-wide reform effort and an enhanced kindergarten literacy program) in place in two of the participating schools associated with gains in reading performance over time for students in this sample?

Based on theories of linguistic interdependence (Cummins, 2001) and the “common underlying proficiency” (Cummins, 1991), we hypothesized that students’ tested reading performance would drop when they transitioned from Spanish (the primary language) to English, but would return to or surpass pre-transition levels. Strong results from a school-wide reform effort at another of the district’s schools in the 4 years prior to the beginning of the present longitudinal study led us to hypothesize that a similar reform, characterized by teacher work groups to establish and evaluate language arts content standards, would contribute to higher levels of performance with students in our study. However, because significant improvements took several years to achieve in the earlier effort (Goldenberg & Gallimore, 1991), we hypothesized that a school-wide reform would not initially affect student performance but would show effects in the intermediate grades. Finally, we hypothesized that a strong early literacy instructional program in kindergarten, characterized by phonics emphasis, academic push, and focused home/school connections, would provide students with an initial boost that would decline over time but would continue to impact reading performance in the intermediate grades and through middle school. The effectiveness of systematic phonics instruction has been demonstrated as early as Chall’s influential study of the “great debate” in reading (Chall, 1967) and reaffirmed in studies since (Adams, 1990; Anderson, Hiebert, Wilkinson, & Scott, 1985). However, recent studies found that the impact of phonics instruction on reading was significantly greater in the early grades than in the later grades (second through sixth grades) (Ehri, 2004).

Method

Sample

Sample recruitment

The longitudinal sample ($N = 121$; 66 boys, 55 girls) was originally recruited in fall of 1989. With the assistance of school officials, we sent out contact letters to the parents of all Spanish-speaking children in 11 different classrooms in four schools in two Los Angeles-area school districts. Of the 296 letters sent home, 252 (85%) were returned by parents who indicated willingness to participate in the study. Parents were then contacted at random by phone. In all, we contacted 154 families to construct a longitudinal cohort of 121 families with a child entering kindergarten. In 15 cases, it was not possible to contact the family after repeated attempts. In another 6 cases we were given incorrect telephone numbers. Six families had moved away by time of the telephone contact. Five families contacted declined to participate after we called and explained the project. One family was omitted because the family had insisted the child be placed in an English-only instructional program at the school. All the other children included in the cohort were scheduled to be placed in Spanish-language-reading instruction at the time the study began.

Sample description

Among the cohort of 121 children when the study began in 1989, 91 lived in Lawson (a pseudonym), an unincorporated area of approximately 1.2 square miles in metropolitan Los Angeles. School enrollment in the Lawson District is over 90% Latino. A second group ($n = 30$) included immigrant Spanish-speaking families residing in a racially mixed neighborhood approximately 25 miles south of Lawson; these children attended school in a large urban district.

The great majority (84%) of the parents in both communities came to the United States from Mexico; the rest were from Central America. When the study began, mothers in the sample averaged 9.6 years (range = 1–34) in the United States; fathers averaged 11.7 (range = 1–53). The average number of years of formal schooling for both mothers and fathers is 7 (range = 0–16). Only 4% of the homes in which the target children reside are headed by widows or single mothers, and one is headed by a widower. In all other homes, two parents are present.

In contrast to their parents, the majority of the children (75%) were born in the United States, 94% in California. Close to 22% of the children were born in Mexico; 3% in Central America. Spanish is the home language of all of the families in the sample. However, in some homes English is used as well. Upon entering kindergarten, none of the students tested fluent in English; however, 24% had some English proficiency.

The great majority of the parents are employed in skilled and unskilled labor jobs. Of the fathers, 45% report being in unskilled labor, 50% are in skilled labor or supervisory positions (e.g., line supervisor in a factory), 1% are in technical or white-collar jobs, and 4% are in managerial positions (e.g., manager of a fast food restaurant). Only 3% of the fathers reported being unemployed in 1989. This percentage rose during the economic recession of the early 1990s (deep and lasting in California), dropping again in the mid-1990s, with 6% of the fathers reported being unemployed in 1995 when the last in-home interview with all families took place. Approximately 43% of the mothers worked outside the home when the study began, with somewhat more being currently employed outside of the home.

Mobility

The mobility of this low-income immigrant population poses challenges for sample retention and gathering of data from the students' schools and teachers. For example, by the end of the first year of the study (the students' kindergarten year), only 62.8% ($n = 76$) of the children were in the same classrooms in which they began the year. An additional 25.6% ($n = 31$) had moved but remained within the same school district. However, 14 students (11.6%) had moved out of the original district, and of these 9 were lost to the project. By the end of elementary school, less than half of the students (44.6%; $n = 54$) remained in their original schools. The students in the sample were enrolled in 32 different schools by the end of elementary school. Most of these schools provided test and teacher rating information for the study; however, each year some data from schools outside of the original participating districts were not made available. The high mobility of our sample is consistent with findings from other studies, which found Latino students more mobile than non-Hispanic White students between Grades 8 and 12 in California and twice as likely as White students to change schools due to adolescent request (Larson & Rumberger, 1998).

Sample retention

Table 1 presents the sample attrition data for the longitudinal sample since the original recruitment in 1989. Over the last 5 years of the current funded project, the retention rate has remained steady. Comparisons of the 91 retained cases with the 30 lost since 1989 indicated no significant differences on various demographic and child variables.

Measures

To gauge students' achievement, we collected data using a variety of measures: standardized tests of reading and math performance, individually administered tests of early literacy performance, individually administered tests of oral English proficiency, and teacher ratings of student performance. For the present study we make use of the early literacy assessments

Table 1

Longitudinal Sample Retention, Fall 1989 to Spring 1998

Year	Number of participants	Percent of participants retained
1989	121	100
1990	112	92.6
1991	106	87.6
1992	103	85.1
1993	97	80.2
1994	91	75.3
1995	90	74.4
1996	92	76.2
1997	92	76.2
1998	91	75.3

(kindergarten) and standardized tests of reading performance (Grades 1 through 8).

From first grade onward, we used standardized test scores normally collected by the schools in the two participating districts. One district used the Spanish Assessment of Basic Education (SABE) in Spanish and the Comprehensive Test of Basic Skills (CTBS) in English. The other district made use of the Comprehensive Test of Basic Skills in Spanish (CTBS-Español) and the Metropolitan Achievement Test (MAT) in English. However, as children progressed through elementary school and middle school, many families moved out of the original participating districts. Standardized test scores (national percentiles in reading and mathematics) continued to be collected for the test(s) used by the different schools and districts; however, collection of test scores for out-of-district students was extremely problematic, and scores for close to half of the out-of-district students were not obtained on a yearly basis. By the end of the students' eighth-grade year, national percentile scores from six different tests were collected (CTBS, California Achievement Test, Individual Test of Academic Skills, Stanford Achievement Test, Iowa Test of Basic Skills, and Texas Assessment of Academic Skills); all of these were in English. National percentile scores were used in the analyses because

these were the scores that were consistently available from the school districts. During the 1993–1994 academic year, one of the original participating districts changed to a different standardized test with local rather than national norms. These test scores, which tended to be higher than national percentile scores, could not be used in our analyses, thus reducing the cases included.

Achievement at the beginning and end of kindergarten was tested using a Spanish-language literacy assessment, the *Prueba de Lectura Inicial* or *Lectura Inicial*, developed for the project because Spanish was the primary language of all participating children. The assessment was administered individually to children over 2 days in two 30-minute sessions. The *Lectura Inicial* asks children to identify letters and their corresponding sounds; read words (real and nonsense); write or attempt to write letters, words, and stories; demonstrate knowledge of important concepts about print; and demonstrate oral comprehension of a story (see Reese et al., 2000). These are important dimensions of early literacy development (Adams, 1990; Clay, 1985; McCormick & Mason, 1989). Raw scores from the test scores at the end of the year on the six subtests of the *Lectura Inicial* were converted to a single factor score; components analysis yielded a single robust factor, with factor loadings of individual measures ranging from .49 to .66 (eigenvalue = 3.26). Some students entered kindergarten unable to recognize any letters or answer questions about a story read aloud to them, and with few notions of what a book is, how it is used, and which part of the book is read. Other students demonstrated proficiency in all areas. By the end of the year, although average performance had increased overall, the range of performance was even greater than had been observed at the beginning of the year. For the study presented here, we used *Lectura Inicial* factor scores (based on raw scores from the end of kindergarten assessments) to categorize the sample into low-, average-, and high-achiever groups.

Data Analysis

For the analyses of individual student achievement trajectories over time, the data were analyzed using a mixed general linear model (MGLM), otherwise known as a random effects or hierarchical regression model. In these models, each subject's scores are treated as an individual sequence, and differences among subjects as well as differences within subjects are recognized. By incorporating group similarities across subjects, it is possible to compensate for missing data points within individual subjects.

The basic assumption is that each child's score at each time point is composed of two indistinguishable parts, one based on predictors and a random error of measurement. In contrast to the usual linear regression model, where the predictor part would be a fixed function, the MGLM assumes that each child has a distinct regression function, and that those functions are selected from a population of regression functions. To reach inference about

the entire population, one studies the properties of this population, and to reach inference about a particular child one estimates the function specific to that child.

Findings

Transition to English

At the time of the study, in order for students to transition from Spanish to English literacy instruction, many districts had a procedure in place to assess student progress in oral English proficiency and in Spanish reading. This procedure followed California Department of Education guidelines and established criteria for transition: oral English fluency based on an approved measure, and assessed Spanish-reading proficiency at approximately the third-grade level. The two districts in which the participating students were originally enrolled used, respectively, the IDEA Proficiency Test and Bilingual Syntax Measure assessments to assess oral English proficiency and used the SABE to assess performance in Spanish reading. Table 2 shows the number of students tested in the spring of each year in English and Spanish on nationally normed standardized tests. In the transitional programs studied, once students transitioned to English they were no longer tested in Spanish. A few students who had been arbitrarily placed in English programs but were at early stages of English proficiency were not tested in either language in Grade 1.

Typically students must demonstrate reading ability in Spanish at approximately low third-grade level in order to begin transition. As is evident in Table 2, in our sample over half of the students had moved into English-reading and English-only instruction by the end of Grade 3; and by the end of Grade 4, close to 84% of the students had made the transition. However, given application of state-required criteria, it is not surprising that a small number of the students did not transition at all during elementary school because they

Table 2

Number of Students Tested in Spanish and English, Grades 1–8 (Spring Only)

	Grade							
	1	2	3	4	5	6	7	8
Spanish	62	60	33	9	6	5	0	0
English	14	18	39	56	49	64	66	61
Total	76	78	72	65	55	69	66	61

did not meet designated levels of proficiency either in oral English or in Spanish reading. On the other hand, in some cases students were automatically placed in English instruction in middle school, regardless of whether or not their assessment indicated sufficiently high proficiency levels to master curricular content in their second language. In addition, a handful of students were placed in English reading as early as kindergarten, either by teacher placement or because the family moved and room was not available in the bilingual program at the receiving school, which is a reflection of the lack of systematic implementation of transitional criteria cross schools within districts.

Figure 1 illustrates mean reading scores by language of instruction. Because a variety of nationally normed standardized tests were used as families moved and children changed districts, we used national percentile scores reported by the schools in our analyses. When mean reading scores are examined over time, a different pattern emerges for each language. In Spanish, reading scores averaged slightly above the 50th percentile and remained high through Grade 2. However, mean scores in Spanish plummeted by late elementary school. This is not because the overall achievement of the group was declining; rather, each year the students who were most proficient in Spanish reading made the transition to English and were no longer tested in Spanish in the transitional programs in the study. By Grades 5 and 6, the only students remaining in the Spanish-reading program were those who, after 6 years in U.S. schools, were not yet proficient in conversational English and/or

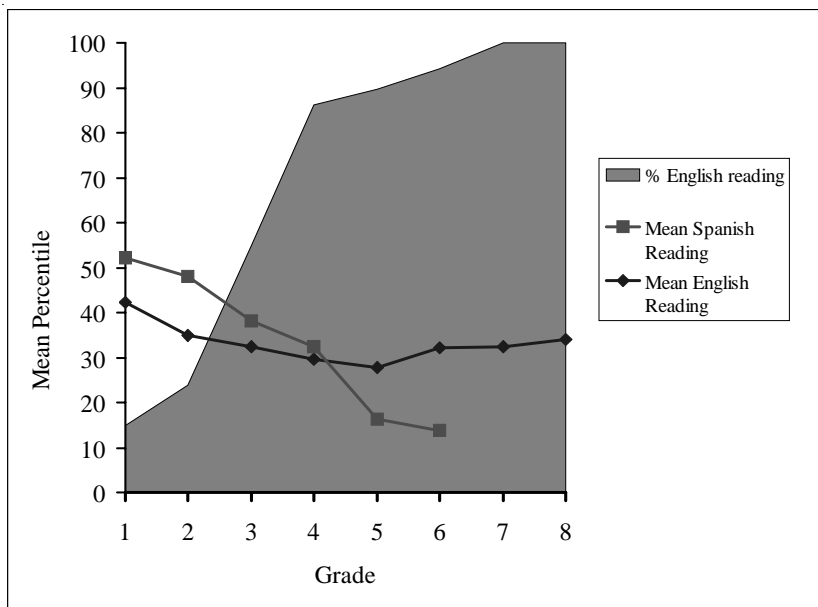


Figure 1. Mean Spanish- and English-reading test scores from first to eighth grade.

were unable to read well enough in Spanish to pass Grade 3-level assessments. On the other hand, the English scores represent those students who, each year, met minimal criteria for transition. Although some students were doing quite well in English reading, the group mean was below .5 standard deviation (*SD*) below the norm and remained at that level throughout elementary and middle school.

Overall, for the total sample, reading performance on standardized tests in Spanish and English combined declined over the years through Grade 4 and then began to improve slightly (see Table 3). The mean on reading national percentile, which was 51.1 in Grade 1, declined sharply to 29.8 in Grade 4, moving up to 34.5 in Grade 8. In general, although mean scores were dropping from year to year while students were receiving instruction in their native language, a sharper drop corresponds to the students' switch from Spanish reading to English. When trajectories are analyzed by grade of transition, all students are observed to plummet when transition occurs, but higher achievers make the transition earlier (typically in Grade 2 or 3) and recover earlier and more quickly. Lower achievers transition later and are slower to recover. Therefore, by Grade 7, the grade of transition to English was negatively correlated with performance ($r = -.48, p < .0001$).

Another way of examining the impact of transition from primary-language reading instruction to second-language instruction is using a MGLM. For this analysis all available scores are used (including fall scores from one district in Grades 1 through 3); however, only students with sufficient data points are included in the analysis. Therefore, numbers of students included vary slightly from those reported in Table 2. In our context, we have assumed that the child's regression function for reading achievement is composed of four parts analogous to a multiple linear regression: (a) an intercept or starting value, (b) a regression slope representing growth in Spanish, (c) a "penalty" representing the transition to English, and (d) another regression slope representing growth in English. For a given child the transition to English may occur at any grade. We used the SAS PROC MIXED program (Littell, Milliken, Stroup, & Wolfinger, 1996) to do the computations and found these estimated

Table 3
Spanish- and English-Reading Trajectories (Means for Sample)

	Grade							
	1	2	3	4	5	6	7	8
Spanish	54.9	51.4	38.9	33.1	17.3	13.8	–	–
English	33.1	35.9	33.3	29.1	28.5	32.1	32.1	34.5
Total	51.1	46.9	35.9	29.8	28.9	32.5	32.9	34.5

population values in percentile achievement scores: intercept = 65.2%, a downward slope in Spanish = -5.8% per year ($p < .0001$), penalty for transition to English (or decline associated with transition to English) = 43.3% ($p < .0001$), and “recovery” slope in English = 1.3% per year ($p < .03$). These values represent the entire group of children; each child has four estimated values to represent his or her own course.

Thus, findings are clear that student performance declined sharply with transition to English, and overall performance of students in transitional programs in English reading did not return to performance levels in Spanish prior to transition.

Differentiated Trajectories

The group means reported in Table 3 mask substantial variation within the sample. To investigate variation in reading achievement trajectories, we categorized the student sample in terms of their literacy and reading development at the end of kindergarten. Using factor scores from the *Lectura Inicial*, we categorized individuals into one of three groups: low achievers ($n = 29$) were considered those who scored .6 or more *SDs* below the mean on the *Lectura Inicial*; average achievers ($n = 28$) scored between .5 *SD* above the mean and .5 *SD* below the mean; and high achievers ($n = 24$) scored .6 *SD* or more above the mean.

Figure 2 presents the mean standardized reading achievement scores from first to eighth grades for the three groups. As Figure 2 reveals, only the group that was scoring well above the mean at the end of kindergarten managed to end middle school reading at a level that approximated where they began. The high group experienced a sharp decline in average scores between Grades 2 and 3, by which time 70% had made the transition from Spanish to English instruction and testing. Another 21% of the high group transitioned to English in fourth grade. Their scores immediately begin to rise after transition but did not reach pre-transition levels until the end of middle school.

On the other hand, there was a marked decline in reading achievement for students who ended kindergarten at or below grade-level reading in Spanish. (average and low groups). These students’ scores began to decline in Spanish prior to transition to English, and continued to decline after transition. By the end of middle school, the early advantage enjoyed by the average group had disappeared, with both average and low groups’ mean performance in English reading falling below the 20th percentile.

A note should be made regarding the numbers of students tested each year on standardized achievement tests. As families moved and students attended schools other than the original cooperating schools, collection of test scores from the school districts became increasingly difficult. In some cases, districts were using assessments that were not nationally normed. In other cases, districts did not respond to requests for test data. Therefore, the

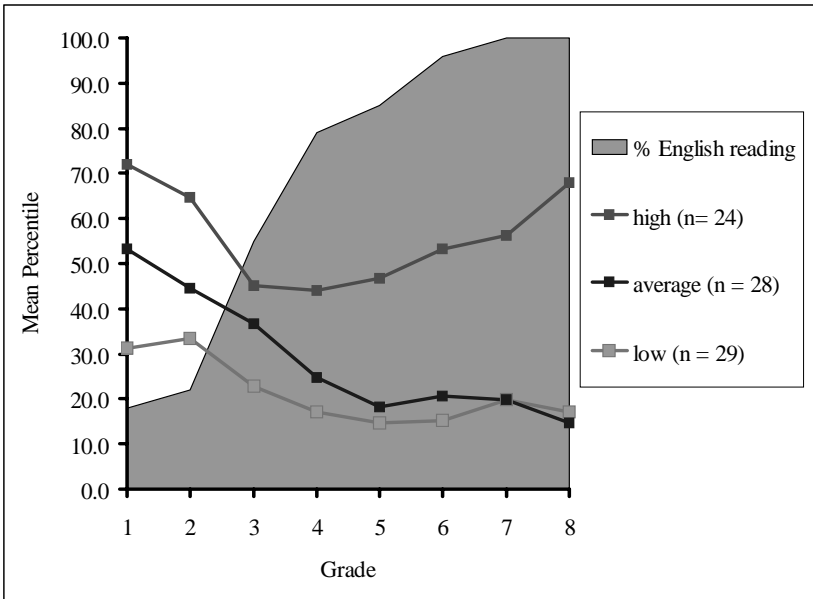


Figure 2. Mean reading achievement for groups categorized by score of post-kindergarten literacy assessment.

Table 4

Mean Reading Achievement for Total Number of Students Tested, Grades 1–8

	Grade							
	1	2	3	4	5	6	7	8
<i>N</i>	81	80	73	59	51	65	64	53
<i>M</i>	51.92	47.31	33.73	28.88	27.90	30.83	32.95	34.50
<i>SD</i>	24.07	24.59	22.37	20.61	22.64	26.55	26.49	37.61

number of students tested refers to the number of students for whom test scores were received. Table 4 presents the available number of cases for each grade level, along with the group means and standard deviations for standardized reading tests obtained from school records.

The percentage of youngsters in English instruction grew steadily, with most children in English by fourth grade. On average, students in the high group transitioned to English earlier (mean grade of transition = 2.58) than the

low (3.36) and average (3.92) groups. In bilingual programs in the districts involved in the study, students were typically transitioned when they achieved a stipulated level of proficiency in Spanish reading and also had conversational ability in English. Thus, students in the high achievement group appear to be advantaged not only in their English-speaking ability in the early grades but also in their reading proficiency in Spanish.

Instructional Interventions

Kindergarten literacy program

Our sample was constructed such that one quarter of the sample was made up of students at an elementary school in which the kindergarten bilingual teachers had developed and implemented a strong early literacy instructional program that had produced gains in student performance in kindergarten. This Spanish literacy program was characterized by an emphasis on phonics instruction, academic push, and focused home-school connections. For example, students received nightly packets of homework, and failure to return homework the next day resulted in a telephone call to the parents. Although some parents in our sample commented that these teachers were too demanding, most were highly satisfied with their children's academic progress and instruction they were receiving in kindergarten.

Families of the children in this program did not differ from others in the sample when compared on dimensions such as length of time in the United States, parents' level of education, and their occupational status. On project-administered tests of initial kindergarten performance on early Spanish literacy skills, overall skill levels were low for all children. For example, the mean number of letter sounds recognized was less than one (range = 0–10). The only subtest which showed a significant difference between the strong kindergarten program group and the others in the sample was on letter names, where the strong program group recognized an average of four letters in comparison to the rest of the group's average of two letters. This suggests that because the administration of individual measures took approximately 1 month to complete the effects of the instructional program focus on letters and literacy were already apparent.

Our findings confirmed what school personnel had suspected, that students who received this intense literacy instruction in kindergarten outperformed their peers in first grade (see Table 5). Moreover, the mean scores for students in this kindergarten cohort, who were dispersed into different classrooms throughout the school (and to different schools as families moved away), show a continued advantage enjoyed by these students over time. This advantage is substantial and significant through Grade 3, after which it drops and is no longer significant. However, the group still outperforms its sample peers by close to 9 percentile points as late as Grade 8. These

Table 5

*Intensive Kindergarten Reading Program(K Program)
Compared With Other Programs*

	Grade							
	1	2	3	4	5	6	7	8
K program	67.5	65.0	46.4	37.5	31.8	34.3	37.4	39.8
All others	43.7	41.1	30.1	25.9	25.9	28.6	29.9	30.9
Difference	23.8	23.9	16.3	12.7	6.9	5.7	7.4	8.9
Significance	.0012	.0002	.0039	ns	ns	ns	ns	ns

findings appear to replicate meta-analyses of National Reading Panel findings that systematic phonics instruction was particularly advantageous in the early grades (Ehri, 2004).

School-wide reform effort

One quarter of our sample was selected from a school in the same district as the school in which the intensive kindergarten literacy program was located. In the fall of 1989, under the leadership of a principal new to the school, they had begun a school-wide reform effort characterized by the creation of teacher work groups at each grade level to establish and evaluate language arts content standards. These work groups met regularly, assisted by a team of university researchers and supported by a principal dedicated to the reform effort. Work groups analyzed student writing, established grade-level expectations for language arts performance, and developed rubrics for assessing student writing based on these expectations.

Although substantial improvement in student reading and writing was documented at this school over a 5-year period (Goldenberg, 2004; Goldenberg & Saunders, 1996), this improvement was not translated into gains on standardized tests of reading for children in our sample. Means for the school reform group were lower than those of other groups throughout middle school, although these differences between the school reform group and all other groups were not significant. It is probable that the students in our sample were "ahead of the curve," moving through the grades just ahead of the grades that received the full impact of school reform efforts, since the efforts did not turn out to be fully implemented during the 1989–1990 school year. For example, language arts assessments showed a jump in achievement on third-grade assessments in 1994, 1 year later than when the students in our sample were in Grade 3 (Goldenberg, 2004).

Discussion

The data we have presented paint a picture of unacceptably low levels of reading performance for many students in our sample; however, by the end of middle school, none of the students had dropped out of school. They were still in the educational system, but it was a system that was increasingly failing them.

In our randomly selected 1989 sample, as many as 33% of youngsters were scoring one or more standard deviations below national norms in reading in English at the end of eighth grade. Yet, there are also indications that these youth were profiting from regular instruction as recently as first or second grade. Such performance levels in early elementary school suggest that this is not a group ordinarily eligible for special education although by the end of middle school they might be considered for placement in many school districts (MacMillan & Speece, 1999). Despite their academic difficulties, by the end of middle school we had no indication of widespread alienation or reaction from these students. In interviews carried out with a sub-sample of students, most still said that they regarded schooling as important, and few had been reported by their teachers to be exhibiting problematic behavior. For the most part, they were engaged with and involved with their families (Gallimore & Reese, 1999).

Findings indicate clearly the “hit” taken by English language learners as they transitioned from primary language to English instruction. The MGLM analysis revealed a substantial 43 percentile point drop in reading achievement associated with transition. However, contrary to predictions, this drop was neither temporary nor was it overcome by the majority of the students in the sample. Although some students recovered quickly and surpassed pre-transition achievement levels, most students did not. Even students who had sufficient family literacy experiences and initial literacy instruction at school to score, on average, slightly above the 50th percentile on standardized tests of reading in Spanish in Grade 1 were not, at least by the end of middle school, able to recuperate and achieve at comparable levels in English.

There are several possible explanations for these findings. One possible explanation, following the linguistic interdependence model, is that the transitional programs in which our children studied did not provide opportunities for prolonged academic development in the primary language. The majority of the students in our sample transitioned in Grade 3, and by the end of Grade 4 all but a handful of students were in English-reading programs. However, close to a quarter of the sample had transitioned by the end of Grade 2. In the transitional programs in which the students were enrolled, formal instruction in Spanish did not continue after transition. It can be argued that the literacy and academic skills that these students had developed in their primary language were not sufficient to support the development of more

complex skills and concepts in English, and students were not provided with the opportunity to continue to develop academic concepts and vocabulary in their primary language once they were transitioned to English.

Another possible explanation of the findings is that the Spanish-reading scores were inflated in comparison to the English scores. Although the CTBS-Español and the SABE tests were both nationally normed, the norming populations for the English standardized reading tests and the Spanish tests are different. Using the example of the English and Spanish versions of the Wechsler Intelligence Scale for Children language test, Valdés and Figueroa (1994) explained how use of samples that were not comparable with respect to SES resulted in differences in raw-score points per scale. More generally, they highlighted the problems inherent in drawing conclusions based on bilingual tests that had either been normed on a national sample from outside of the United States or from a regional sample within the United States, and questioned the validity of such measures. Therefore, with respect to children in our sample, a comparison of English and Spanish scores is certainly problematic, and to expect a “recuperation” in English back to tested levels in Spanish might be unrealistic.

A third possibility, the one implied by the criticisms of bilingual education and the decision to end bilingual programs by voter initiative in California (Proposition 227), is that students may not be receiving sufficient English language development as part of their bilingual programs. Therefore, when they move from Spanish to English, their oral English proficiency is not solid enough to support grade-level academic learning and literacy. Whereas transitional bilingual programs typically have taught literacy sequentially, beginning in the primary language, a recent study of existing evidence found bilingual approaches to be superior to English-only approaches with the strongest effects for English language learners when they received literacy instruction in both English and the primary language from the onset (Slavin & Cheung, 2003).

A final possible explanation is that the program of instruction received by the majority of students in the sample, regardless of language of instruction, was ineffective and poorly implemented. Although our study did not include program evaluation as its goal, we documented cases in which the program was not implemented as planned. For example, a few students were “flip flopped” between Spanish and English reading from year to year while others were moved into English because the teacher felt they were “mature enough” to progress in English. Our results certainly indicate that, for the majority of students in our sample, the transitional program did not result in high levels of performance. The children who prospered in the programs studied were students who entered school with literacy resources fostered outside of elementary school. For the second-generation children in our sample, literacy performance in middle school was predicted by greater exposure to literacy

activities (in Spanish) at home and greater exposure to English in preschool as well as through extended residence in this country (Reese et al., 2000). Those students who entered school with these assets—greater familiarity with Spanish literacy developed through home activities and greater proficiency in English—were the students who left kindergarten doing well in reading and for whom the transitional bilingual program was successful. For students who began school with more modest assets and ended kindergarten with somewhat lower reading performance, the instructional program did not promote and sustain academic progress either in Spanish or in English. In their recommendations for improving schools for language-minority students, August and Hakuta (1997) stressed the importance of moving beyond evaluations of which type of program was best and focusing instead on program quality and on the instructional components necessary for student success in local communities that may vary in terms of goals, demographics, and resources.

Our data cannot test these competing views. It is likely that all, in some combination, play a role in explaining our findings. Our data do speak to the urgency of finding ways to build on children's early linguistic and experiential knowledge in ways that allow all to achieve to their full potential.

References

- Adams, M. (1990). *Beginning to read: Thinking and learning about print*. Cambridge, MA: MIT Press.
- Anderson, R., Hiebert, E., Wilkinson, I., & Scott, J. (1985). *Becoming a nation of readers*. Champaign, IL: Center for the Study of Reading.
- August, D. & Hakuta, K. (1997). *Improving schooling for language-minority children: A research agenda*. Washington, DC: National Academy Press.
- Baker, C. (1996). *Foundations of bilingual education and bilingualism* (2nd ed.). Philadelphia: Multilingual Matters.
- Clay, M. (1985). *The early detection of reading difficulties* (3rd ed.). Portsmouth, NH: Heinemann.
- Cummins, J. (1991). The role of primary language development in promoting educational success for language minority students. In California Department of Education (Ed.), *Schooling and language minority students: A theoretical framework* (pp. 3–49). Los Angeles: Evaluation, Dissemination and Assessment Center, California State University, Los Angeles.
- Cummins, J. (2001). *Negotiating identities: Education for empowerment in a diverse society* (2nd ed.). Los Angeles: California Association for Bilingual Education.

- Ehri, L. (2004). Teaching phonemic awareness and phonics: An explanation of the National Reading Panel meta-analyses. In P. McCardle & V. Chhabra (Eds.), *The voice of evidence in reading research* (pp. 153–186). Baltimore: Paul H. Brookes.
- Gallimore, R., & Reese, L. (1999). Mexican immigrants in urban California: Forging adaptations from familiar and new cultural sources. In M. C. Foblets & P. C. Lin (Eds.) *Culture, ethnicity, and migration* (pp. 245–264). Leuven, Belgium: Acco.
- Goldenberg, C. (2004). *Successful school change: Creating settings to improve teaching and learning*. New York: Teachers College Press.
- Goldenberg, C., & Gallimore, R. (1991). Local knowledge, research knowledge, and educational change: A case study of early Spanish reading improvement. *Educational Researcher*, 20(8), 2–14.
- Goldenberg, C., Gallimore, R., Reese, L. J., & Garnier, H. (2000). Cause or effect? A longitudinal study of immigrant Latinos' educational aspirations and expectations and their children's school performance. *American Educational Research Journal*, 38(3), 547–582.
- Goldenberg, C., & Saunders, W. (1996). Using standards and assessments to promote excellence and equity among diverse students. In R. Blum & J. Arter (Eds.), *Who's learning what? A resource guide for student performance assessment in the context of school restructuring* (pp. 1–4). Alexandria, VA: Association for Supervision and Curriculum Development.
- Larson, K., & Rumberger, R. (1998, February). *School mobility and California Latino adolescents*. Paper presented at the annual conference of the UC Linguistic Minority Research Institute, Santa Cruz, CA.
- Littell, R. C., Milliken, G. A., Stroup, W. W., & Wolfinger, R. D. (1996). *SAS system for mixed models*. Cary, NC: SAS Institute.
- MacMillan, D. L., & Speece, D. L. (1999). Utility of current diagnostic categories for research and practice. In R. Gallimore, L. Bernheimer, D. MacMillan, D. Speece, & S. Vaughn (Eds.), *Developmental perspectives on high incidence handicapping conditions: Papers in honor of Barbara K. Keogh* (pp. 111–133). Mahwah, NJ: Erlbaum & Associates.
- McCormick, C., & Mason, J. (1989). Fostering reading for Head Start children with little books. In J. Allen & J.M. Mason (Eds.), *Risk makers, risk takers, risk breakers: Reducing the risks for young literacy learners* (pp. 154–177). Portsmouth, NH: Heinemann.
- Ramírez, D. (1992). Executive summary. *Bilingual Research Journal*, 16(1–2), 1–61.

- Reese, L., Garnier, H., Gallimore, R., & Goldenberg, C. (2000). A longitudinal analysis of the ecocultural antecedents of emergent Spanish literacy and subsequent English reading achievement of Spanish-speaking students. *American Educational Research Journal*, 37(3), 633–662.
- Siantz, M. L. (1997). *A global profile of the immigrant child*. Washington, DC: American Academy of Nursing.
- Slavin, R., & Cheung, A. (2003). *Effective reading programs for English language learners: A best-evidence synthesis* (Report No. 66). Baltimore: Johns Hopkins University, Center for Research on the Education of Students Placed at Risk.
- Suárez-Orozco, M. M., & Suárez-Orozco, C. (1997). *Longitudinal factors in immigrant adaptation*. Washington, DC: Division of Social, Behavioral, and Economic Research, National Science Foundation, Award No. 9706278.
- Thomas, W., & Collier, V. (1997). *School effectiveness for language minority students*. Washington, DC: National Clearinghouse for Bilingual Education.
- Thomas, W., & Collier, V. (2002). *A national study of school effectiveness for language minority students' long-term academic achievement* (Final Report: Project 1.1). Santa Cruz, CA: Center for Research on Education, Diversity, and Excellence.
- Valdés, G., & Figueroa, R. (1994). *Bilingualism and testing: A special case of bias*. Westport, CN: Ablex.

Acknowledgments

This research was supported by grants from the National Institute of Child Health and Human Development and the Spencer Foundation, with additional support provided by the Center for Culture and Health. Our thanks go to the parents, children, and school personnel who made this work possible and special thanks to Claude Goldenberg, our collaborator on the project from which this study is drawn.