

A National Study of School Effectiveness for Language Minority Students’ Long-Term Academic Achievement

Findings from Two Rural Research Sites in the Northeast U.S.

The Regional Social Context

Two school districts that participated in our study—Madawaska School Department and School Administrative District #24—are located in northern Maine, along the U.S. border with Canada, in the St. John Valley. Each small town and the surrounding area that these school districts serve has a total population of 4-5,000 people. Student mobility in the schools is relatively low, since most remain in the community until graduation from high school. Over 90 percent of the students of these rural school districts are of Franco-American/Acadian heritage. While U.S. Census data indicate that French is the mother tongue of 97 percent of the residents of the towns served by these two school districts, French has actually been in strong decline in this region over the past half-century. The “power and status” language is English. An ambivalent view of bilingualism that developed during the 20th century among community members, parents, school personnel, and students, gradually led to a negative self-image among adult members of the francophone community, accompanied by lower school achievement of francophone students.

Cummins (2000, pp. 41-42) discusses the situation of minority francophone students in Canada, which parallels the experience of francophone students on the U.S. side of the border.

Sociological analyses indicate that these francophone communities experience life in ways similar to subordinate minorities in other parts of the world. Power and identity issues come to the fore in institutions such as school, where the process of integration and assimilation creates a slow destruction of ethnolinguistic identity. Marginalized groups often experience ambivalence and

insecurity in relation to their bilingual/bicultural identity, resulting in low achievement in school and underemployment in the workplace. Devaluation of the particular variety of the French language spoken in the border region further creates ambivalence, resulting in francophones' hostility toward the majority language group, shame for their own culture, and poor academic self-perception and performance.

Before the school innovation that we analyzed for our study was introduced in the two school districts of this border region, the francophone students were experiencing a high degree of subtractive bilingualism, with students gradually losing their oral French as they acquired oral and written English. Losing one's first language (L1) as second language (L2) is acquired generally leads to lower achievement in school; whereas additive bilingualism—acquiring L2 at no cost to L1—generally leads to high achievement in school (Lambert, 1975). Despite the francophone origins of the students, very few students used the French language to any significant degree in the home or community, due to the high level of linguistic assimilation within the community because past generations did not have the opportunity to be schooled in French (Landry & Allard, 1992). In fact, the previous generation, parents of the current students, were punished for speaking French in school. A school board member explained, “We’ve been brought up for a long time to see French as a street language, not worthy to be taught. We never learned to read it and write it ... It was not important enough to have in school” (Hoose, 1996).

Given this sociocultural context, the two school districts participating in this study chose to try to reverse the patterns of lower academic achievement among some of the francophone students, when compared to the achievement levels of anglophone and other bilingual francophone students in the same schools. To accomplish this goal, they developed an enrichment bilingual program for all students who chose to participate in the French-English classes, beginning with Grades K-2, expanding it to the upper primary grades, and eventually to middle school and high school

levels. We were able to collect longitudinal data on students receiving bilingual schooling in Grades K-7 during the five years of this study, since the new program was implemented at more than one grade level with each successive year of program implementation.

This enrichment bilingual program was labeled a two-way model by the program implementers. However, while the classes in these school districts include a few anglophone students, the large majority (90%) are students of Franco-American heritage. Thus, for purposes of our research study, we are classifying this as a one-way program as we have defined the model—there is one ethnolinguistic group being schooled through their two community languages (rather than two language groups receiving schooling through their two languages). Both one-way and two-way are for all practical purposes the same type of bilingual program, in that both are integrated, mainstream, enrichment models, designed for all students who choose to attend. But in a context such as northern Maine where almost all students are of the same ethnolinguistic background, the classes include fewer students who are proficient in one of the two languages—in this case, French. In the program, these francophone students have less access to peer models who speak, read, and write French proficiently, to stimulate the development of natural French language acquisition. They are therefore more likely to develop higher proficiency in English (their stronger community language) than in French.

Contrasting Patterns In This Region

Higher English proficiency. This particular context for one-way bilingual education is somewhat different from any of our other research sites, so it provides an interesting contrast. First, these francophone students were reasonably proficient in English at the start of the project. In fact, on self-rating scales administered to students in Grade 2, students stated that English is the language most often used in all social domains, including the home. Students tended to rate their ability to speak English as high and their ability to speak French as moderately low. On

standardized tests in English given in 1993-1994, before the program started, students designated as the target group because they were less proficient in academic English initially scored as a group at the 40th NCE (31st percentile) on the subtests of reading and mathematics of the California Test of Basic Skills. In contrast, the comparison group of high-achieving francophone students in these two school districts reached the 58th NCE on the reading subtest and the 60th NCE on the mathematics subtest by the end of elementary school. Thus, one goal of the project was to raise the academic achievement of the lower-achieving francophones to at least comparable levels with their counterparts who were doing quite well in school. This starting point for the lower-achieving francophone students is significantly higher than that of any of our other subjects in other research sites. But the ultimate goal is gap reduction, so in our findings, we examine the students' starting point and follow their progress across time from their initial achievement level to their end attainment as of the final year of data collection.

Heritage language and culture restoration for purposes of economic development. A second distinction from our other research sites is this project's focus on restoration of a minority language, as a means to promoting higher academic achievement, eventually leading to economic revitalization of the region. This one-way bilingual program can be categorized as focused on linguistic and cultural revitalization, incorporating both language restoration of the students' heritage language as a major goal, as well as bicultural identity formation, for the building of higher self-esteem among the francophone students. In this context, the schools are working on promotion of the heritage language of their region, even though the francophone community mostly speaks the majority language, English.

This goal was initially difficult to promote in the community. French had been denigrated for most of the parents' lives. Furthermore, the particular French variety spoken in the region was perceived negatively. Yet the region is quite isolated from the English-speaking United States. It

is 350 miles to Portland, Maine, and 200 miles to any other substantial urban area of the U.S.; and the harsh winters frequently make these large distances impassable. It is a quick drive to French-speaking Canada, the main social and economic stimulus for businesses in Northern Maine. Proficient adult bilinguals could assist the economic growth of the region. A school board member explains, “The world’s getting smaller. We have NAFTA (the North American Free Trade Agreement), and many of our businesses cater to Canadian tourists and the French-speaking elderly population.”

Furthermore, these communities had already experimented with a transitional bilingual program. The implicit goal of the transitional program was to replace the minority language with the majority language—classes in French served as a transition to English, the majority, high status language. But the transitional program did not significantly raise students’ test scores. Thus the initiators of the new one-way enrichment bilingual program chose to make the goals more explicit—to foster true additive bilingualism and biliteracy in both French and English for all students, as well as to foster knowledge of, and pride in, the local culture.

The school implementers intended to have a significant influence on improving the prevailing community attitudes toward the Franco-American/Acadian language and cultural heritage. Also they wanted to improve the language and literacy skills and the academic achievement of francophone school-age children. The bilingual school staff expected initial resistance among parents, since teaching through students’ heritage language is often perceived as delaying and limiting the children’s access to the language of power, English. Thus, in the project’s design, they provided for bicultural community involvement, extensive training and support for school personnel, and the development of substantial bilingual/bicultural community resources to create meaningful and challenging academic curricula through the community’s two languages. Dramatic changes have already occurred during the five years of implementation of the program to date, as

confirmed by school board members and school administrators. “We’re seeing a big change in community attitudes toward the French language. The kids just love singing French songs and speaking and writing French. It’s like it’s become a new fad, for children and adults,” stated one school board member.

Status of French and local language varieties. A third unusual aspect of this one-way bilingual project is its focus on developing a language less valued by the local community, but that in other contexts would be considered a prestigious language. French is considered a high-status language around the world, an important language of diplomacy, a language commonly chosen to be studied in schools throughout the world. Of the slightly less than 200 sovereign states of the world, 120 states have adopted English (54 countries), French (33 countries), Spanish, or Arabic as their official language (Baker & Prys Jones, 1998, p. 346). But the French language variety spoken in northern Maine and across the border in Canada has lower status than the “Parisian” variety. This regional variation still has many similarities to its counterpart in France. Regional vocabulary can be perceived positively if examined from the point of view of additive bilingualism. Thus teachers in this program support the bidialectalism that students develop in French, adding to their oral knowledge from the local variety the standard oral and written forms of French. The result is students who are metalinguistically aware and increasingly proficient in academic and business uses of French across multiple regional contexts. During the English portion of the academic day, students get continuing high-level development of oral and written English for academic purposes, which also prepares them for their use of the two languages in the adult world.

Established ethnolinguistic minority. A fourth distinguishing feature of the social context of these two research sites is that the target group for the planned school innovation is not an immigrant group. They are an established ethnolinguistic group with a history that goes back to

the first settlements by Europeans in the Americas. While the francophones in this region of the U.S. would be classified as a minority by the U.S. government, they are the majority (over 97 percent) for this region. In fact, the francophones of the northeast U.S. have experienced some sociolinguistic and socioeconomic patterns similar to that of the indigenous Spanish-speaking peoples of the southwest U.S. Schools in the southwest also prohibited the use of Spanish in school during parts of the past century, and the regional varieties of Spanish spoken in the southwest are generally perceived as lower status than the standard varieties of Spanish spoken in each Latin American country and Spain. In rural and urban areas along the U.S.-Mexican border, there is considerable poverty. In northern Maine, the economy is quite depressed on the U.S. side of the border, but flourishing on the Canadian side. So there are some parallel issues to be examined in the schooling provided for historically established ethnolinguistic groups living near a U.S. border.

Implementation of the School Innovation

Balance of the two languages of instruction. The school program chosen to improve francophone student achievement was started in the school year 1996-1997. Planning for the new project took place during the year prior to implementation. Initially, the planners wanted to implement a 90-10 model, similar to that practiced in Canadian and California bilingual immersion programs, with 90 percent of the instruction in the lower grades (K-1) in the minority language (French), gradually increasing the percentage of instruction in English until by around 4th grade, 50 percent of the instruction would be in each language. However, in the first year of the project, teachers' implementation practices remained closer to a 50-50 balance of the two languages. During the second year of the project, teachers tried to emphasize more French in the early grades, with an approximate ratio of 60-40. The planners pushed hard to increase French instruction to 70 percent for Grades K-3 by the third year of implementation. Resource teachers reported that by Years 3 and 4, the goal of 70 percent French instruction had been achieved in some K-3 classrooms, but not all. Overall, as actually implemented, this could be classified as a

50-50 model, for Grades K-8.

The pressure to teach more in English is strong. Politically, it was difficult to promote the 90-10 model to some of the principals and parents. Generally, teachers were very positive about the program goals but also realistic about the level of bilingualism that could be attained with this amount of French instruction, and they would have preferred increased time in French, given that English is the stronger language of these communities. Research on bilingual immersion programs demonstrates quite clearly that more teaching in the minority language, provided in the 90-10 model, contributes to higher level skills in that language with no apparent loss to English (Baker & Prys Jones, 1998; Cummins, 2000; Genesee, 1987). But the strong social pressure in the U.S. to learn and use only English overwhelms the ideal teaching situation for proficient bilingualism to be developed.

Grade levels for implementation of program. The program began in 1996-97 in Grades K-2. Since oral usage of the French dialect of the region was common among students of all ages, and their proficiency level in English was relatively high at the start of the program, it was not necessary to introduce the project one grade level at a time. This implementation practice is common in two-way bilingual programs when each group of students is just beginning acquisition of one of the languages of instruction. Using the hands-on curriculum of kindergarten and first grade, students acquire their new language through curricular tasks, so that by the time the students reach second grade and beyond, they have had enough exposure to the new language to be able to do second grade work in their second language. Instead, in these schools the project was implemented by blocks of grade levels to give the program planners time to collect and develop the materials and resources for each block. Thus in the second year, the program was extended to Grades K-5 (K-6 in one school district), and in the third year, the program reached Grades K-7. By the fifth year of the project (2000-2001), there were sufficient numbers of

eighth graders who had received their schooling through the two languages to offer coursework in French as well as English to that grade level, completing K-8 implementation. Plans are in place for continuation at high school level, by offering some courses for credit in French.

Teacher credentials. All 15 bilingual teachers in the project have certification for the age group or subject area that they are teaching, 5 have bilingual teacher certification in addition to certification to teach in mainstream classes, and one-third have completed masters degrees. One-third of the 120 teachers in the two school districts are proficient in both oral and written French and English, and three-fourths of the teachers use French for oral purposes. Six of the 12 administrators of the schools use French orally and 3 administrators have some written French proficiency. Thus the teaching and administrative staff are supportive of bilingualism and a sufficient number of teachers proficient in French were already in place when the project began to teach thematic units in the content areas through French. Extensive staff development was provided by the project to expand teachers' repertoire in French in all content areas. The staff are mostly of Franco-American heritage; thus students and staff share a common cultural heritage.

Students. When the program began, there was a total of approximately 900 students in school district #1 and 600 in school district #2. Among the francophone students with lower achievement scores who were designated most in need of the bilingual program, 154 of those in school district #1 chose to participate in the bilingual classes and 104 of those in school district #2 chose to participate in the bilingual classes. Other students doing well academically were allowed to enroll too. In all, approximately one-third of the total student body participated in bilingual immersion classes, most being of Franco-American heritage.

Teaching style. The 15 bilingual teachers were regularly visited by bilingual resource teachers who assisted with staff development and planning needs. As a part of our data collection efforts,

the resource teachers completed a survey regarding classroom practices among the bilingual teachers. The issue of translation was especially noteworthy in the findings from this survey. Seven of the 15 teachers stated that they translate for students and that students are allowed to use both languages as needed. The remaining eight teachers maintained the research ideal that in a bilingual class, the teacher separates the use of the two languages and does not translate. The program evaluator mentioned that in his classroom observations he found more reliance on translation than should be allowed for high proficiency in French to develop. Unfortunately, formal measures of French written proficiency were not administered in the project, so there is no way to examine the students' level of academic French development. However, students gained substantially on the oral French measures informally assessed by teachers.

Almost all of the bilingual teachers described their use of thematic lessons, cooperative learning, hands-on instructional materials with lots of visuals and manipulatives, use of microcomputers, multicultural literature, journal writing, and authentic assessment as important aspects of their teaching. All of the teachers reported that they connect the curriculum to the students' experiences, and 13 agreed that they incorporate bicultural knowledge into the curriculum. Six teachers stated that they use the community and parents' knowledge regularly as a resource for student learning. Nine of the bilingual teachers integrate art, music, and drama into their thematic units that develop language arts, mathematics, science, and social studies knowledge.

Revitalization of Franco-American/Acadian culture. To stimulate the connections to community and cultural heritage, the program directors and resource staff have arranged many varied activities that have helped students and staff and the community revisit and celebrate their cultural heritage, for raising self-esteem in the community, with the ultimate goal being economic revitalization of the region. These activities have included high school-elementary school partnerships for students to read books in French together; parents' participation in storytelling

and book reading in French; guest speakers with backgrounds in Acadian and Native American history and culture; field trips to nearby Acadian and Quebec cities and villages; francophone state senators' and university professors' presentations to school board members, administrators, and teachers; numerous courses and training workshops for teachers on Franco-American/Acadian history and culture and French language and literature for each age group; positive newspaper articles and TV news programs disseminating accomplishments of the bilingual program; developing a resource collection of videos for families on Franco-American/Acadian culture; intensive summer French language institutes for teachers and students in France; and co-sponsoring community events with the Acadian Archives, a state cultural performance center, a state French theatre group, the National Park Service, local chambers of commerce, state universities, the Franco-American Center, and associations connected with French-English bilingual immersion programs in Canada. Parent organizations help sponsor some of these events and the parents of the immersion program print regular issues of a newsletter to keep the community informed and to encourage new parents to enroll their children in the program.

Development of French proficiency. Since the program conducts no formal assessments of the children's developing French proficiency, a few informal measures will be mentioned. One was a self-evaluation by second grade students of their oral communicative proficiency in French, conducted by staff during the third year of the program. A score of 1 indicates great difficulty in communication and a score of 5 indicates that communication is effortless and natural. Students rated themselves at a 3.8-3.9 level when asking for a telephone number or describing members of family. For expressing personal feelings, explaining rules of a game, convincing someone to get involved, and explaining the contents of a course, they rated their communicative skills in French in the mid-range (2.81-3.22). Relating events in a movie was slightly harder (2.66), placing an order for a meal still harder (2.28), discussing important personal problems more difficult (2.19),

and discussing point of view on political system was the most difficult of all to do in French (1.66), as would be natural for a child of this age group.

Teachers regularly conducted informal assessments of students' French proficiency development, but these were not available to the researchers. A parent's words best describe the feelings expressed by both the teachers and parents in the program: "It's a fantastic program! We have no regrets of putting her in the program because she has picked up so much. Kristin reads and writes French very well. She's at the top of her class. Even in math concepts in French, she just picks it up like a sponge ... and it has not bothered her English. It's amazing, just wicked amazing! It's unbelievable the difference; she has picked up tremendously. It's been a great challenge for her and myself. The teachers are excellent."

Summary of Social Context and Implementation Findings

An evaluator of the program summarizes the linguistic and cultural goals of this program, contrasting this with the reality of its social context:

This dual-language immersion program constitutes an attempt by a linguistic minority to revitalize its culture and language. Unfortunately, this is not an easy task because the 'reversing language shift' process is filled with challenges and obstacles that few linguistic minorities surmount (Fishman, 1989, 1990). Ideally, a community with such a strong demographic capital (close to 97% of French ethnolinguistic origins) should have maintained a strong vitality. However, the community did not have sufficient 'institutional completeness' (Breton, 1964) to develop a community life capable of fostering intergenerational transfer of the group's mother tongue. Such a task was manageable until relatively recently, even without having access to French schooling, because of a certain degree of isolation,

proximity to other Francophones in the neighboring provinces of New Brunswick and Quebec in Canada, and an economic lifestyle that did not require a high degree of education or mobility. Today, the growing urbanization, the highly influential effects of the mass media, the more pervasive contacts with the dominant language and other factors contribute to an increasing difficulty to maintain intact one's culture in a minority context. Furthermore, until recently, the use of French at school was prohibited and punished. Many parents have been left stigmatized by their negative experiences and low self-esteem. The bilingual education program is an attempt to counter negative attitudes and to promote pride in the community. The challenge is great because a large proportion of the children attending school today have not acquired French as a first language despite the fact that their parents and many of the same generation community members still understand and speak the language.

The socializing process in French is now left almost exclusively to the school. The school has the arduous task of counterbalancing the strong forces of assimilation now present in society. Such a task would be a momentous challenge even if the whole school was a unilingual refuge in French. The challenge is even greater within a bilingual school where the dominant language is clearly English. (Landry, 1997, pp. 15-16)

Landry proposes continuing to create within the schools "community life" where grandparents, parents, teachers, and children can meet and participate in meaningful social and cultural activities that will promote pride and legitimize the affirmative expression of a bilingual/bicultural identity. Increasing the percentage of time devoted to teaching in French while maintaining the high standards in the teaching of the English language will also contribute to higher levels of additive

bilingualism. Ironically, for this research study, our only quantitative measures are student performance on English academic achievement tests, but these measures do demonstrate the powerful stimulus of receiving enrichment schooling through two languages at no cost to English.

Results in Student Academic Achievement

Cross-sectional analyses. Our analyses include cross-sectional, quasi-longitudinal, and longitudinal views of the data, through descriptive, analysis of variance, and multiple regression analyses. The cross-sectional analyses are presented in Tables A-1 through A-4. These first four data displays do not follow precisely the same group of students across time and therefore they are labeled cross-sectional, although there is very little student mobility in these two school districts, so many of the students who started bilingual immersion classes remained in those classes throughout the years of this study. These tables examine all students attending bilingual immersion classes in 1997, 1998, 1999, and 2000, and compare their performance to all students in the English mainstream classes, as measured by their normal curve equivalent (NCE) scores on the Terra Nova, a standardized, national norm-referenced test, on the reading, language, and mathematics subtests.

As can be seen in Table A-1, in 1997, with the new program implemented in Grades K-2, the bilingual immersion students outperformed the monolingually schooled students at the end of the first year of the program by at least 5 NCEs in reading, language, and mathematics. Also, the bilingually schooled students scored well above grade level, except for the second graders' math performance, which was slightly above the 50th NCE/percentile and thus on grade level. It was expected that the bilingual immersion students would score lower than monolingually schooled students in the first year of the program, as is found in most research studies on bilingual education. However, these students did quite well in Grades 1-3, during the first two years of implementation. In Table A-2, the fourth and sixth graders in 1998 may have scored lower

because this was the first year of implementation of bilingual instruction in these grades, as the program was expanded from K-2 to K-6 in this year.

Since both school districts' typical performance on these tests is above the 50th NCE/percentile, it is remarkable that by 1999 (see Tables A-3 and A-4), the bilingual immersion students were outscoring the monolingually schooled students at all grade levels, sometimes by very significant amounts. (In this research, the criterion for statistical significance is set at $\alpha < .05$, unless otherwise noted.) As seen in Table A-3, on the 1999 Terra Nova, the bilingual immersion students outperformed the monolingually schooled students at every grade level by 4-17 NCEs, except second grade math, fifth grade reading, and fifth grade language arts, which were 3 NCEs lower, and not a significant difference, statistically or practically. On the 2000 Terra Nova (Table A-4), the bilingual immersion students again outperformed the English mainstream students, except in second grade language arts (3 NCEs lower—not significant), second grade math (6 NCEs lower—the only lower difference with significance) and seventh grade math (1 NCE lower—not significant). By the end of seventh grade, after three years of bilingual schooling in French and English, the bilingual immersion students were scoring at the 57th NCE (63rd percentile) in reading, the 60th NCE (68th percentile) in English language arts, and the 55th NCE (60th percentile) in math. In the last year of testing for this study (2000), the bilingual immersion students consistently scored at or above the 55th NCE (60th percentile) in every subject and every grade level, except for second grade math at the 48th NCE (47th percentile).

Quasi-longitudinal analyses. Figures A-1 and A-2 and Table A-5 present quasi-longitudinal analyses, meaning that the same groups were followed across time (e.g. those with one year of bilingual schooling, those with two years, etc.), but there is no available pre-test measure to compare to the post-test scores in this particular analysis. These figures clearly show the impact that the bilingual immersion program had on the students who were designated as most in need of

the program. Franco-American students who were achieving less well in the two school districts were identified as a target group for the bilingual program, when it was first conceived. At the end of the school year 1993-94, this group of lower-achieving students was scoring at the 40th NCE in reading and mathematics on the California Test of Basic Skills (CTBS). In the figures and tables this group is designated as “Former LEP” (limited-English-proficient). This group of former LEPs is made up of two groups—those who received the bilingual immersion program and those who received English mainstream instruction. The other group, labeled “Non-LEP,” is also of Franco-American heritage, but they were achieving above grade level when they began the bilingual immersion program, at the 58th NCE in reading and the 60th NCE in math on the CTBS in 1994.

As can be seen in Figure A-1 and Table A-5, the students designated former LEP, who tested at the 40th NCE (31st percentile) in 1994 on the reading subtest of the CTBS, gained with each additional year in the bilingual immersion program on the reading subtest of the Terra Nova, from the 47th NCE (1 year of bilingual schooling), to the 51st NCE (2 years), the 56th NCE (3 years), and reaching the 62nd NCE (71st percentile) after 4 years of bilingual schooling. For the four-year group, this is a steady and significant gain of 22 NCEs in reading achievement—both statistically and practically significant, and a difference equivalent to a full national standard deviation. The former LEP comparison group, taken from the same group of low-achieving francophone students, were those whose parents chose for their children not to be schooled in the bilingual classes. This group of former LEP students, schooled all in English, ended at the 48th NCE in reading in the year 2000, a significant but smaller gain of 8 NCEs. Thus, the bilingual immersion students gained an average of 5.5 NCEs per year, while the English-instructed students gained an average of 2 NCEs per year.

In the language arts subtest, former LEP students made similar gains with each additional year of

schooling in the bilingual immersion program, from the 46th NCE to the 59th and 58th NCEs, finishing at the 61st NCE (70th percentile) after four years of bilingual schooling. Their four-year comparison group being schooled all in English was at the 50th NCE on the language arts subtest in 2000. In the math subtest, former LEP students in bilingual classes achieved at the 48th NCE after one year, the 53rd after two years, down to the 47th after three years, and at the 59th NCE (66th percentile) after four years in bilingual immersion. In 2000 their comparison group schooled all in English was at the 50th NCE in math.

The students in the bilingual immersion program who were designated as non-LEP, those achieving above grade level when the program started, also benefitted greatly from their schooling through two languages. In 2000, the non-LEP students in the all-English curriculum scored at the 53rd, 56th, and 56th NCEs respectively in reading, language, and math; while their non-LEP counterparts who had received four years in the bilingual French-English curriculum scored on the same three subjects at the 61st, 61st and 59th NCEs. Overall, these results are very significant, statistically and practically, strongly favoring the bilingual immersion program.

Longitudinal analyses. The remaining descriptive analyses, examining the data from a longitudinal perspective, are presented in Figures A-3 through A-5 and Tables A-6 through A-11. These analyses consist only of those students with pre-tests in 1997 and post-tests in 2000. Note that the pre-tests were administered after one year of program operation; thus, the pre-post gains do not include effects from the first year of the program. These figures and tables again demonstrate the high achievement of the bilingual immersion students, in comparison to the monolingually schooled students. The different achievement levels are most evident when examining the students who were former LEPs. On the 1997 reading measure, after one year of the program, the former LEPs in bilingual classes were 9-12 NCEs above the achievement level of the former LEPs in the all-English classes at each grade level, except fifth grade at 3 NCEs above.

On the 2000 reading subtest at all grade levels these same bilingual students were 6-10 NCEs above their comparison group attending all-English classes. Similar patterns occurred in the language arts and math scores of the two former LEP groups, with the bilingual students outperforming the all-English students by 1-12 NCEs, except for three cases where both groups scored equally high.

Three cautionary notes should be mentioned for interpreting the gain scores in Tables A-6, A-8, and A-10. When the numbers of students are broken down by grade level and by number of years in the program, the numbers in each group are sometimes too low for the test score mean to be reliable. As a criterion, we suggest that when the number (N) for a group is less than 10, the average scores should not be considered. Since the group sizes in our longitudinal analyses are quite low, we consider the quasi-longitudinal analyses to be more valid for decision-making purposes.

The second point regarding interpreting gain scores as measured in NCEs (normal curve equivalents) or percentiles is that we want to make sure the readers understand the difference between these types of scores and the scores given by teachers on classroom tests or scaled standard scores on norm-referenced and criterion-referenced tests. The NCE is not a cumulative score of the total number of points correctly answered on the test. Instead, both NCEs and percentiles are rankings of how well a group of students did in relation to the typical performance of all students in the U.S. However, the amount of student achievement in percentiles changes across the range of percentile values, because of the shape of the normal distribution. NCEs correct this problem and thus may be considered as equal-interval or corrected percentiles. The 50th NCE/percentile means that 50 percent of the students in the U.S. at that grade level scored below that level and 50 percent scored above that level of performance. If a group of students stays at the same NCE level of achievement from one school year to the next, making a zero NCE

gain, that means that they have made one whole year's progress. If they were on grade level (around the 50th percentile) the past year, a zero NCE gain means they stayed on grade level over one year's time. For students to gain in NCEs from one year to the next, it means that they have made more than one year's progress during the year. Scoring 2-3 NCEs above or below the previous year's performance is generally within the standard error of the mean, so that amount of change is not considered significant. However, a difference of 4 NCEs or more can be considered significant, since this difference is equivalent to an effect size of 0.2 or more.

The third issue also involves interpretation of the gain scores, in the third column of each of Tables A-6 through A-11. In an analysis where students are initially scoring very low (e.g. the 10th NCE, a common starting point for LEP students when first tested on norm-referenced tests in English), gain scores are meaningful and important, since the overall goal is gap closure when comparing to typical native-English speakers scoring at the 50th NCE. However, it is worth noting that pre-test scores may be "falsely low" if obtained before the LEP student has sufficiently mastered enough English to take a test administered in English. Thus, in a short-term study comparing beginning-of-year pre-test scores to end-of-year post-test scores, the resulting gains may be too large because the pre-test scores were unreliable. This problem tends to disappear as LEP students learn enough English to enable them to effectively take the test on the same basis as native-English-speaking students. Thus, a long-term study that follows LEP students' annual gains for several years avoids this problem while short-term studies may suffer from it.

Groups that initially are scoring low need to make more than one year's progress each year for several years in a row to eventually close the gap. But in these two school districts, the students who were not former LEPs were scoring at or above the 50th NCE at the start of this program. Those groups of students in these tables who are initially scoring at a high level—e.g. the 55th-60th

NCE—generally reach a ceiling. In other words, high achieving students do not typically continue to achieve higher with every year of school, but at some point, their scores reach an above-average range and stay there. Thus the gain scores are less meaningful for students already scoring above grade level. Some columns show a net loss of 2-3 NCEs between the pretest in 1997 and the posttest in 2000. But if the students' scores were already high on the pretest, that net loss is not considered significant (because it is less than two-tenths of a national standard deviation, expressed in NCEs).

Tables A-7 (reading), A-9 (language arts), and A-11 (math) present the longitudinal test results combined across all grades, so that the number of students in each group is sufficiently large to be able to make more reliable comparisons. These analyses show patterns similar to the cross-sectional results, continuing to confirm that the bilingually schooled students, both in 1997 and in 2000, have clearly outperformed their monolingually schooled comparison groups.

In 1997, on reading, language arts, and math, the former LEP students in bilingual classes scored at the 55th, 60th, and 59th NCEs—8, 7, and 6 NCEs higher than their comparison group, the former LEP students in all-English classes. These two groups of former LEPs started at the 40th NCE in reading and math achievement on the norm-referenced test administered the year before the program began. On the 2000 test examining the same three subjects, the former LEP students in bilingual classes scored at the 56th, 59th, and 55th NCEs—8, 8, and 3 NCEs higher than their former LEP comparison group attending all-English classes. This testing includes Grades 4-7, so that even with the increased cognitive complexity of the middle school tests, the bilingually schooled students have been able to maintain their above-grade performance.

The non-LEP comparison groups followed a pattern of achievement similar to that of the LEP comparison groups. In 1997, non-LEP students schooled bilingually scored at the 62nd, 65th, and

65th NCEs—4, 4, and 3 NCEs higher than their monolingually schooled counterparts. In 2000, the bilingually schooled students scored at the 60th, 63rd, and 63rd NCEs—5, 4, and 5 NCEs higher than their comparison group in all-English classes. These two groups combined started at the 58th and 60th NCEs in reading and math respectively in the year before the program began.

Thus, among both LEPs and non-LEPs, those who were schooled bilingually outscored those schooled monolingually after both one year and four years of the bilingual program. All of these results—cross-sectional, quasi-longitudinal, and longitudinal—dramatically demonstrate that students schooled through two languages outperform those schooled through one language. These bilingually schooled students have also acquired French at no cost to their English achievement.

Repeated measures analyses of variance. In addition to tests of practical significance of findings that rely on effect sizes, we have conducted tests of statistical significance. We rely mainly on practical significance, given that the power of statistical tests is very much influenced by sample sizes, leading to Type II errors (failure to find statistical significance) when sample sizes are too small, and to Type I errors (false finding of statistical significance) when sample sizes are too large. Thus, since group sizes are routinely small in the Maine data, we believe that practical significance of observed differences among groups, expressed in conservative terms of fractions of a national standard deviation of 21.06 NCEs, is a better guide for policy making than is statistical significance of differences.

In conducting statistical tests of significance where both pre-test and post-test scores are available, and where the measurement scales are the same for both, we have chosen repeated-measures ANOVA to assess the observed differences both between groups and within groups. The repeated-measures factor is the pre-post achievement test measure (e.g., Total Reading,

Total Math, Total Language) and the independent variables are Program Experience (yes vs. no), LEP Status (LEP vs. non-LEP), and Grade in 2000. Since pre-test and post-test scores are from the same students in a longitudinal study, these scores are correlated, and repeated-measures ANOVA makes use of this to reduce the error term in the F-test, and thus to increase the power of the statistical test. In effect, repeated-measures ANOVA is an extreme form of blocking, since we are blocking on each subject, thus completely removing from the error term all variability among subjects because of individual differences. In repeated-measures ANOVA, the independent variables are assessed not as a main effect, but as an interaction between the independent variable and the pre-post factor. Finally, in all repeated-measures tests described below, the alpha-level for statistical significance is set at .05.

Tables A-12 through A-17 provide the results of the repeated measures ANOVA tests where pre-test and post-test scores on Total Reading achievement (Tables A-12 and A-13), Total Language achievement (Tables A-14 and A-15) and Total Math achievement (Tables A-16 and A-17) are examined in succession. In each pair of tables, the students were nested within the independent variables Program Experience, student LEP Status, and Grade as of the year 2000 in the first table, and within Program Experience and LEP Status in the second table.

In Tables A-12 and A-13, the factor for pre-post Reading between program years 2 and 5 is not significant, and the interactions between pre-post and each independent variable are also not significant, indicating that there is no significant difference between pre- and post-test NCE scores overall or among the groups of the independent variables. Since “no significant difference” in this case means a pre-post gain of zero NCEs, we interpret this to mean that all students have maintained their relative positions in the norm group over time during program years 2-5. Thus, they have made one-year’s-progress-in-one-year’s-time during each year of the program. The tests of between-subjects effects indicate that the students defined by variables LEP Status and

Program Experience were significantly different at both pre-test and at the post-test. Table A-7 provides the pre-test and post-test means for these groups and specifies the direction of these differences; non-LEP scores are higher than LEP scores and the scores of students in the program are higher than the scores of students not in the program. However, it is well worth noting that these scores are measured over the 1997-2000 period only, since the first formal student testing in the program did not occur until Spring 1997. Thus, these 1997-2000 analyses do not reflect program gains made during the first year of program operation in 1996-97, since there was no formal pre-test in Spring 1996. In other words, these 1997-2000 comparisons may well underestimate the program's true effect because any achievement gains made during 1996-97, the program's first year, are not included.

In Tables A-14 and A-15, the pre-post factor of interest is Total Language achievement as measured in Spring 1997 and in Spring 2000. Here, students who were in grade 4 in 2000 (and grade 1 in 1997) show a significant drop in average NCE score of more than 6 NCEs. There are also decreases in the mean NCEs for the grades 5, 6, and 7 in year 2000, but these are smaller, verging on non-significance. Interestingly, it was the non-LEP students, not the LEP students, who contributed much of the observed pre-post decrease in mean NCEs in Total Language. This was true for both the non-LEP students in the program and those not a part of the program. Apparently, the non-LEPs' already high NCE scores were subject to a ceiling effect that did not affect the LEP students as much over time. There were no significant differences among the groups of the between-subjects variables LEP Status and Grade in 2000, but there was a significant difference between program participants and non-participants, with program participants scoring slightly higher at both 1997 and 2000 testings.

In Total Math achievement, as shown in Tables A-16 and A-17, there was a statistically significant decrease between 1997 and 2000 in student math scores. As shown in table A-10, this

decrease occurred among the groups in grades 4 through 7 in Year 2000. This finding is supported by the significant interaction between the Pre-post factor and Grade as shown in Table A-16. Thus, students' NCE scores tended to decrease as the 2000 grade approaches the middle school years, indicating an across-the-board increase in the difficulty of the math test items for all groups as the grades proceed from Grade 4 to Grade 7. There were no significant interactions between Pre-post and Program Experience or between Pre-post and LEP Status, indicating stable trends for these groups between 1997 and 2000. The tests of between-subjects effects indicate a significant difference between LEPs and non-LEPs in math achievement, with non-LEPs having higher scores in both 1997 and 2000.

Stepwise regression analyses. In Table A-18, we used hierarchical stepwise regression to assess the potential impact of two variables on the former LEP students' reading achievement test scores in the year 2000. The procedure is described by Cohen and Cohen (1975).

Each potential predictor is entered into the regression equation first, thus maximizing its opportunity to produce an incremental increase in multiple R squared (R^2). Then each potential predictor is entered into the regression equation last, and the resulting increment in R squared is used to estimate its unique impact on achievement after variance of the first predictor has been accounted for. In this way, we are able to arrive at an estimate of the unique effect of each predictor on reading achievement, and an estimate of their shared effect on achievement.

When the variables "Socioeconomic Status" (SES) as measured by free or reduced lunch, and "Years in Bilingual Classes" are each introduced first, R squared increases from zero by .06 (SES) and .085 (Years in Bilingual Classes) respectively, indicating that the number of years attending bilingual classes exerts a stronger effect on eventual reading achievement in 2000 than former LEP students' socioeconomic status.

When each variable is entered last, the unique effects of each variable (as indicated by the increment in R squared) are estimated at R^2 change=.037 ($F=3.673$, $df=1,86$, $p<.06$) for SES and R^2 change=.06 ($F=5.9$, $df=1,86$, $p<.02$) for Years in Bilingual Classes. The variance shared between these two variables is estimated at .025. This indicates that SES does not exert a statistically significant impact on eventual student reading achievement ($\beta = .05$) but that Years in Bilingual Classes is a statistically significant determinant of eventual reading achievement. In addition, since shared variance could be attributed to either predictor, the impact of Years in Bilingual Classes could be as high as .085 ($F=8.06$, $df=1,87$, $p<.01$). This is worth noting because the number of years that students attend bilingual classes is a “changeable” variable, in that it can be influenced by policy decisions to adopt effective programs for LEP students, whereas students’ socioeconomic status is much more difficult (if not impossible) for the school district to change. This means that decisions to adopt one-way and two-way dual language programs for LEPs might influence as much as 8.5% of eventual LEP reading achievement.

As shown elsewhere in this report, we have observed in other districts that the impact of the program developed for LEP students is substantially higher than seen here when less effective program alternatives (e.g., ESL pullout) are compared to more effective program alternatives (one-way and two-way dual language education—also called bilingual immersion and developmental bilingual education). In all school districts, the effects of socioeconomic status on student achievement are modified (reduced) when more effective programs are adopted. This offers strong evidence that the negative effects of socioeconomic status on student achievement can be mostly “reversed” or overcome when LEP students are provided with effective bilingual program alternatives.

Conclusions

The data from these two rural school districts in northern Maine demonstrate the high levels of student achievement that is possible when students are schooled through two languages. The heritage language of this community, French, has been in strong decline in this region over the past half-century. The power and status language is English. Yet with the commitment of the school administrators and school boards of these two school districts, those families who have chosen for their children to be schooled in both French and English are experiencing dramatic renewal of their heritage language at no cost to their children's English achievement. Overall, those students schooled bilingually are outperforming those schooled monolingually.

The students who were designated as limited in English proficiency before the program began had a significant command of English, in comparison to similarly classified students in other school districts in the U.S. As a group, just before the program began, these students scored at the 40th NCE (31st percentile) in English reading and mathematics. In other states with large numbers of LEP students, often the 40th percentile (45th NCE) is the level at which students are reclassified as fluent in English and ready for the mainstream. Thus these former LEP students were close to the level of reclassification as fluent English speakers ready for the mainstream when they began the program. But during the first four years in the program, they went from the 40th NCE to the 62nd NCE in reading achievement across the curriculum (see Figure A-1). That is dramatic achievement gain for this particular group. Clearly, these former LEP students have benefitted significantly from their schooling in both French and English. They have made significant gains in their academic achievement in English, and at the same time, they have acquired proficiency in their heritage language, French.

Both school districts are achieving at a significantly high level for all groups of students. The comparison groups being schooled monolingually through English are doing quite well, staying above the 50th percentile, even when the academic difficulty increases in the secondary years.

But the students being schooled through their two heritage languages, French and English, are achieving at higher levels than their monolingually schooled peers, and they are adding French to their knowledge base. The community goal with this bilingual program is to produce more student graduates who are academically proficient in both languages of the community, for economic revitalization of the region. Anecdotal stories among school board members and administrators affirm that students and families have significantly benefitted from this school program, through higher self-esteem among former low-achieving francophones who are now high achievers along with their student peers, and through greater pride in the use of French in the community. English is not diminishing in influence—it remains a strong part of the economy. But now these communities have developed the potential for graduating adult bilinguals who will be able to use their two languages in the workplace to stimulate future economic growth as they serve this U.S.-Canadian bilingual region.