English Learners in North Carolina, 2009
Wayne P. Thomas, Professor Emeritus, George Mason University
Virginia P. Collier, Professor Emerita, George Mason University

Introduction

Since the mid-1990s, it has become clear that students whose first language is not English are not faring well in U.S. schools (Collier & Thomas, 2009; Goldenberg, 2008; Lindholm-Leary, 2001; Thomas & Collier, 1997, 2002). In addition, several large research meta-analyses (Greene, 1997; Rolstad, Mahoney & Glass, 2005; Slavin & Cheung, 2003; Willig, 1985) have conclusively shown that traditional school programs for English learners\(^1\) that use only English to help these students master the curriculum and learn their second language (English) have inferior outcomes when compared to programs that capitalize on the students’ first language to help them master their second language, and also master the curriculum in their second language.

In response to these studies, U.S. educators have taken several actions. First, they have tried to improve the educational outcomes of English as a Second Language (ESL) programs for the benefit of English learners. In particular, these improved ESL programs have addressed the developmental needs of English learners in the cognitive, linguistic, and academic domains that might allow these students to close the second half of their achievement gap with native English speakers that is typically unclosed by traditional ESL programs. One of these improvements is called SIOP teacher training (Sheltered Instruction, which focuses on strategies for teaching the second language through all curricular content areas).

Second, educators have worked to improve the outcomes of traditional bilingual education. These programs are mostly transitional bilingual programs first introduced in the 1960s and ‘70s, that typically provide 2-4 years of education aimed at “making the transition” to English by using the students’ first language to assist the learning process. Students in these programs typically close slightly more than half of their achievement gap (faring better than in traditional or improved ESL programs) but do not reach all the way to grade-level achievement in English after exiting the special program.

Third, many educators have focused on meeting the needs of English learners with an advanced form of English-learner education called dual language education. This program type, which serves English learners well, but is also appropriate and advantageous for all students in a school including native-English speakers, has strong theoretical and data-based support in the research literature. It has also been available since the 1960s, but in scattered locations across the U.S., and has become much more popular since the mid-1990s. Since then, several major research studies have documented its power to produce long-term gap closure for English learners, for other language-minority students\(^2\), and even to improve the performance of native-English speakers (Collier & Thomas, 2009; Lindholm-Leary, 2001; Thomas & Collier, 1997, 2002).

\(^1\) We are using the term “English learners” in this introductory section, because we prefer this term to the federal government terminology “limited-English-proficient (LEP)” students, which is also the current term used by the state of North Carolina. The term “limited-English-proficient” is considered very offensive by many students and their parents; thus we are introducing in this report the term widely used in the field “English language learners (ELLs)” and the shorter form “English learners.” However, in our research findings section, we will refer to these students as LEPs, to be consistent with the terminology of the state of North Carolina.

\(^2\) The term “language-minority” is a federal government term, referring to U.S. students who come from a home where a language other than English is spoken (perhaps only by grandparents, but the ancestral language is still an important part of the students’ lives). They may include recently arriving immigrants as well as eighth-generation citizens of the U.S. Language majority students in the U.S. grow up in a home where the majority language, English, is the main or only language spoken.
The North Carolina Context

Like many U.S. states, North Carolina has experienced a tremendous increase in students whose first language is not English. Nationwide, 65% were born in the U.S. to immigrant parents (Education Week, January 8, 2009) and thus are full U.S. citizens, by provision of the U.S. Constitution, and have a citizen’s full rights to equal educational opportunity, as defined by U.S. Supreme Court decisions, by federal and state constitutions, and by federal and state legislation.

While North Carolina has responded to the needs of English learners with increasingly well-developed English-as-a-Second-Language (ESL) programs, preliminary indications are that graduates of these programs still exhibit substantial achievement gaps when compared to students whose first language or home language is English. And so, North Carolina educators and decision-makers have become more interested in the possibilities afforded by dual language programs, as an across-the-board reform that has the potential for becoming the preferred education program for all students in the mainstream classroom in schools and school districts that are interested in this option.

Dual Language Programs

There are several types of these programs, but students in each type attend schools where the curriculum is presented in English for part of the instructional time, and in a second language (e.g., Spanish, Chinese, Japanese, French) for the other curricular portion (at least half of the time, if not more). These programs also intensively focus on English learners’ cognitive, academic, and linguistic developmental needs. The long-term instructional goal is full educational parity with native-English speakers. In other words, educators who develop these programs are working to provide English learners with a comparable mastery of both English and the curriculum to that experienced by native-English speakers in U.S. schools, while also providing for the native-speaker-like mastery of a second language for all participating students. A major distinction between dual language education and foreign language teaching is that the full school curriculum is taught through the two languages, rather than only teaching the new language as a subject for a small portion of the curricular day. In dual language education, strict separation of the two languages is a non-negotiable component, so that students acquire deep proficiency in each language over several years of curricular experience through both languages (Rogers, 2009).

Unlike traditional ESL and transitional bilingual programs that are intended only for English learners, dual language programs are intended for all students, including English learners, as well as students from a language-minority background who are not classified by the schools as English learners, and also for native-English speakers. There are one-way programs that provide instruction in two languages for one language group (e.g., Spanish speakers). There are also integrated, two-way dual language programs that provide instruction in two languages for all groups, working together. In effect, these programs provide the full mainstream curriculum to all students, just as the mainstream curriculum currently does, but provides it in two languages rather than one. Since available research indicates that all three groups in dual language programs perform better than comparable students who do not participate, dual language programs represent an across-the-board improvement over traditional education programs for these three groups (English learners, language minorities, and native-English speakers), and a potential education reform than can improve the outcomes of current U.S. education for all groups.

Racial, Ethnic, and Language Groups

In this document, we frequently refer to the observed achievement of various racial and ethnic groups in North Carolina. This raises a dilemma of which group labels to use as we differentiate one group from
another in our descriptions. We wish to focus the attention of policy makers, educators, and members of the various racial, ethnic, and language groups on our findings. We also wish to use group descriptions and labels that will cause no offense to anyone. Accordingly, we have chosen to be primarily guided by the U.S. Census categories, by the categories established in North Carolina state databases that we have analyzed, and by terminology adopted by the U.S. Department of Education in describing the various groups whose achievement is analyzed in this study.

For racial groups, we refer to Caucasians (Whites), African-Americans (Blacks), Asian-Americans (Asians), American Indians (Native Americans), and use a Multiracial category for those who claim more than one racial group in their ancestry. In contrast, Hispanics (Latinos) represent an ethnic group that can appear in any of the racial groups.

In the case of language groups, we refer to Language Minority students (those who represent a racial/ethnic group from a language background other than English in the U.S.). We also refer to English Learners (Limited English Proficient students or LEPs) -- these are Language Minority students who have been determined by testing to require school support services to receive a “meaningful education” under federal legal guidelines and rules. Finally, we refer to native English speakers, and to those who are Language Minority but not English Learners (LM-not-LEPs).

As a general rule in this report, we use a more formal label (e.g., African-Americans, Asian-Americans) in the text where space is not an issue but we also usually employ a shorter label (e.g., Blacks, Asians) in the tables and graphs where space is at a premium. We hope that these labels will lead to descriptive accuracy in the text, and will increase visual effectiveness in the charts and tables. At the same time, we also hope that no one will be distracted from the issues at hand by these labels and that no one will feel slighted by their use.

**Gap Closure and English Learners**

While the typically large achievement gap between African-American (Black) and Caucasian (White) students has been investigated since the advent of education program evaluation in the late 1960s, accurate documentation of the even larger gap between English learners and native-English speakers is more recent. Since the U.S. Supreme Court decision *Lau v. Nichols* in 1974 requires a “meaningful education” for English learners and outlaws the practice of educating English learners in the mainstream classroom with regular classroom teachers not trained to meet their special needs, English learners have typically received either one of several forms of ESL or one of several different forms of bilingual schooling.

Since the 1970s, researchers have focused on trying to find statistically significant differences in achievement outcomes between various English learner programs (e.g., traditional ESL vs. traditional transitional bilingual education). However, as Thomas and Collier (1997, 2002) demonstrated, the real concern is that both of these alternatives close only about half of the English learners’ achievement gap with native-English speakers in the schools. Thus, a focus on the traditional research question of “which group is better?” tends to focus attention away from the fact that neither of these alternatives fully closes the English learners’ achievement gap. Instead, research must also focus on degree of gap closure as the most appropriate criterion for program success. Thomas and Collier (1997, 2002) insisted that the most meaningful way to evaluate program success for English learners is to focus on the research question of “How much gap closure is associated with each program type?” instead of focusing on “How statistically significant are the differences between program outcomes?” and ignoring gap closure.
A Decision-making Orientation

This research report is intended to provide decision-makers with information and analyses that decision-makers most need to improve and refine the available program alternatives for English learners. These include questions that decision-makers have asked themselves, and questions that researchers have asked on their behalf, based on prevailing theories that purport to describe how, and how well, programs of interest work in the “real world” of the schools. Thus, the analyses, findings, and conclusions presented here are decision-oriented and are intended to provide what decision-makers should know, and what they need to know, to best evaluate, refine, and improve the performance of English learners in the schools. By necessity, this is an interactive process where analyses are performed.

In contrast, much available research is oriented toward conclusions, without regard to how well these conclusions meet the needs of decision-makers. In response, decision-makers either ignore research that answers questions they didn’t ask, or they ignore all research and make decisions anyway, based on factors such as their personal and professional experience.

Significant Differences

Traditionally, significant differences among group scores have been determined using statistical significance tests. However, it is well known that statistical significance can be “driven” by sample size. In particular, large Ns can create statistical tests that are “over-powered” and tend to find significant differences that are too small for “real world” use or too small upon which to take action in school settings. On the other hand, sample Ns that are too small cause the statistical test to be “under-powered” and also tend to cause violation of the assumptions upon which the statistical tests are based, thus rendering the statistical tests useless or even completely misleading. In studies like this one, we encounter a very wide variation of group sizes, from tens to thousands, causing a number of problems with traditional tests of statistical significance.

In order to avoid some of these problems of falsely finding significance (Type I error) or falsely finding no difference between groups (Type II error) in this report, we use the concepts of “practical significance” and an “actionable difference”. Practical significance is defined in terms of effect size, or the percentage of a combined-group standard deviation that is equivalent to the observed difference between groups. Common guidelines call for differences between groups to be at least equivalent to .30 (or about one-third) of a combined-group standard deviation to be considered to have decision-making relevance and to have pragmatic utility in the world of school-based education.

In other words, an observed difference between one group and another should not be considered “real” unless it is at least as large as one-third of a standard deviation for the combined groups. And so, we adopt the convention that a difference that is (a) larger than .30 of the standard deviation of the combined groups and (b) has group sizes or Ns of at least 25, is an “actionable difference,” or a difference that decision-makers can reliably take action upon. This is a somewhat conservative approach, statistically speaking, and we make it even more conservative by mostly using standard deviations of entire Grades of students (e.g., Grade 3, Grade 4) rather the standard deviation of the combined-groups whose between-group differences are being examined. Thus, our findings may somewhat understate the importance of differences that are found significant, in the hopes that this approach will provide decision-makers with assessments and findings that are robust enough to support important educational decisions about matters that affect children, teachers, and parents.
Program Evaluation Questions and Concerns

Five Stages of Analysis and the Research Questions for Each Stage

We typically plan our collaborative work with education entities such as the North Carolina Department of Public Instruction (NC DPI) through five major stages over a period of several years. However, because only one year of state test data is presently in usable form, and because NC DPI staff was not able to provide all of the requested data needed for full completion of these stages, the present report only addresses stage 1 and stage 3. However, in order to provide an overview of the entire multi-year evaluation process, of which this study is a part, in the following pages we are summarizing the full five stages and their intents, evaluative questions, and required data (as excerpted and adapted from Thomas & Collier, 1997, 2002). Please see the following section, entitled “Overview of Program Evaluation Data and Plan of This Study” for the subset of research and evaluation activities conducted in this study and discussed in this report.

This five-stage process initially examines the effects of past programs for English learners, and conducts a needs assessment to determine the size of the achievement gap between English learners and native-English-speaking students of the school districts. We allow the programs of interest to mature to the point that they can be evaluated both feasibly and validly. In this way, we evaluate programs as fully-implemented rather than as planned or as they are developing and being refined.

These five stages provide a useful template for our collaborative research with NC DPI staff. Each of the five stages addresses a different set of research questions and requires data on different variables to be collected and analyzed. As such, the stages generally guide but do not determine our work. As circumstances and preferences differ among decision makers, we modify and customize our procedures to better address local characteristics and needs. However, we continue to address the overall concerns of each stage to the greatest degree possible. This flexibility avoids the “one-size-fits-all” problem in which a research study may sacrifice ecological validity in the interest of achieving a “standardized” research design.

These early stages reflect the program evaluation perspective that appropriate educational inquiry should focus initially on program development, and on improving program processes so that programs are validly and faithfully implemented according to plans and specifications before outcomes-oriented evaluation occurs. Also, the initial stages identify and focus attention on theoretically-based factors that should enable eventual program success in eliminating the achievement gap between native-English speakers and English learners. Only after these matters have been accomplished is it appropriate to engage in summative evaluation for purposes of program comparison of fully developed and mature programs, or in inferential “scientific research” that attempts to control for the effects of extraneous variables.

Stage 1: A focused needs assessment

In Stage 1 analyses, we examine the difference in long-term achievement levels between three mutually exclusive groups: (1) former English learners who have received local programs designed for English learners; (2) language-minority students who were not classified as English learners and were not in programs specially designed for English learners; (3) and non-language-minority native-English speakers. Naming this comparison the Thomas-Collier Test of Equal Educational Opportunity, we have required each of the participating school districts in our past research to examine this comparison as a condition of working with us. The Thomas-Collier Test establishes whether a school district’s programs for English learners are presently allowing these students to reach long-term achievement parity with non-English-learners in the district. It also provides an opportunity for districts to disaggregate the test scores for two
frequently combined categories of language-minority students: those who have been classified as English learners and are eligible for services, and those who are not.

This is a useful disaggregation for policy purposes because we have noted in the past that many school districts have “hidden” (intentionally or unintentionally) their English learners’ large achievement gap by reporting together the achievement of English learners and those who are members of language-minority groups but are considered proficient in English. Also, districts have focused only on the short-term achievement of these groups, ignoring the fact that achievement gaps continue to develop over time. Finally, districts may focus on achievement gaps of one group while de-emphasizing the performance lags of other groups of interest.

Stage 1 analyses address this issue by comparing the achievement of language-minority English learners served by special ESL/bilingual programs, language-minority students designated proficient in English who are not served by special programs, and non-language-minority native-English speakers. In addition, these major groups may be further divided into sub-groups based on ethnicity, socioeconomic status as measured by receiving free or reduced lunch, or other conditions, such as those students whose parents refuse English learner services offered by the schools. In this way, a clearer and more accurate picture of the impact of local programs on all student groups emerges.

Using the results of these analyses, the district can decide not to address these issues and drop out of our collaborative agreement, or decide to address these issues by continuing on to the successive stages of our joint research. Since we began our collaborative work in 1986, no participating school district has chosen to ignore the findings of Stage 1 analyses and drop out of our collaborative evaluation work. In addition, no school district has yet completely “passed” the Thomas-Collier Test initially, prior to our extensive joint work on developing the district’s programs for English learners and other “at-risk” groups.

Stage 2: English learners = academic achievement gains by length of residence in the U.S. and age on arrival

In Stage 2 analyses, we focus only on English learners and examine their achievement gains over the past 3-5 years. We break down their achievement gains by the students’ length of residency in the U.S. (in the case of immigrants) or number of years of exposure to English. In addition, we break down achievement gains by student age upon entry into ESL/bilingual programs or, for immigrants, by their age on arrival. We have found in prior research that English learners’ abilities to close the achievement gap differ greatly depending on whether they are participating in an ESL or bilingual program or have left the special program and entered the regular instructional program. Since these prior findings imply that length of special program, as well as program quality, are both important factors in closing the large achievement gap, we devote Stage 2 to a thorough investigation of these matters. Thus, these analyses serve to confirm the findings of Stage 1 and to further explore how the observed achievement gap has developed in the school district. Neither Stage 1 nor Stage 2 examines the particular programs that English learners received, but Stages 3, 4, and 5 do.

Stage 3: Achievement gap closure by program type

In Stage 3 analyses, we examine the degree of achievement gap closure that characterizes each program type offered for English learners by the school district during the past five years or more. Each program is described by its average rate of gap closure or achievement gain but no attempt is made to control for extraneous variables at this point because only average achievement gain per year is being examined over a relatively short period of several years. The research question of interest here is, “Looking at trend data
in a time-series fashion, what has been the average progress of students in each program type, measured as average gain and degree of achievement gap closure? Programs in which English learners reach and maintain grade-level achievement with each additional year of schooling, are deemed more effective overall than programs with little or no demonstrated gap closure.

Stage 3 analyses serve several very important functions. First, they provide school district decision-makers with interim, formative information on student achievement that allows a “time-series” comparison of the effectiveness of their various program offerings for English learners over the past several years. This is a pragmatic response to the political needs of school boards, superintendents, and program administrators to have in-progress interim results from their efforts to design better programs for English learners. These groups are simply unable and unwilling to wait for many years to know whether their efforts to improve programs for English learners are productive or not.

Second, Stage 3 analyses provide useful information to the districts as to which of their past programs for English learners have demonstrably closed the achievement gap and which have not. This information can be very enlightening to both administrators and teachers who may be personally convinced of the efficacy of one program type or another, but have never actually examined how student graduates of their preferred program really perform in long-term school achievement, as assessed by the same tests given to native-English speakers, on-grade-level and in English. Such information is made even more useful when conclusions and findings can be confirmed across multiple groups and contexts. Stage 4 addresses these issues of generalizability.

**Stage 4: Increasing longitudinal sample size**

In Stage 4, we add more cohorts of students and use cross-validation and re-sampling techniques to improve generalizability of findings. We add as many years of student data and as many longitudinal cohorts of the same students followed over time as are available and reasonable to add, to further increase sample sizes. This stage addresses the problem of student attrition caused by students leaving the school district. In addition, adding more student cohorts and groups provides opportunities to compare findings to those from the initial study, to investigate separately any groups whose findings differ significantly from those of similar groups, looking for possible moderator or “hidden” variables whose effects on local student achievement had not been previously recognized.

Also, in some instances of Stage 4 work, we use re-sampling techniques (e.g., the bootstrap), a set of statistical methods that yield valid population parameter estimates from local sample statistics to achieve more generalizable estimates of the long-term impact of special programs for English learners. This adds robustness to our later conclusions about which English learner programs have been most successful in closing the achievement gap and keeping it closed as students move through their K-12 school years.

Only after the work of Stages 1-4 has been completed is it appropriate to take up questions of summative long-term program effectiveness in Stage 5. This is true because it typically takes years to achieve: (1) full development of each program to its design specifications; (2) full training of the professional staff to understand each program’s instructional features and to deliver these features as designed; (3) full development of an adequate data-collection system in the school district that will allow ongoing analyses of instructionally important variables and student characteristics over time, and not be limited to the typical 1-2 year data collection time frames in which most school districts operate.
Stage 5: Comparing program outcomes while controlling for extraneous variables

Finally, in Stage 5 of our analyses, we turn to the research question, “Which program is better, when extraneous variables (e.g., initial differences between groups) are controlled?” In this stage, we use more sophisticated analysis techniques such as repeated-measures ANOVA, multiple regression analyses, and controls for extraneous variables (both experimental and statistical) when comparing program outcomes.

These analyses are appropriate only after two conditions have been met. First, the programs for English learners must have “matured” past the point of initial program installation. Second, the programs must have reached a point of full implementation by the school district that is faithful to the specifications and theoretical design features of each of the programs. Otherwise, level and quality of implementation is confounded with program type, resulting in the comparison of poorly implemented programs of one type with well implemented programs of another type. In order to arrive at valid between-program comparisons, all programs must be meeting their full theoretical potential in terms of implementation, at least to the point that is pragmatically possible within the context of good administrative support and well-trained teachers.

In Stage 3, we collect information on degree and quality of program implementation in each school. We accomplish this by means of surveys directed to each classroom teacher, by interviews with instructional coordinators who observe instruction in the schools for each program, and analyzing any data collected by the school district on how instruction is carried out in each school. These data are added to the data collection system and provide possible variables for use in Stage 5.

Collaborative interpretation of data analyses

When the data analyses from Stages 1-3 are completed, we consult with school district staff in collaborative interpretation of the results. Sometimes this leads to the decision to collect additional data, or to reanalyze the data, focusing on new or revised research questions of local interest. We then continue with Stages 4 and 5. This process can re-cycle back to earlier stages if this is deemed useful, continues to succeeding stages when appropriate, and over time, provides data-based guidance to administrators and teachers for constructive changes in school policies and programs, collaboratively agreed upon. If the school districts wish to continue in this cyclical reform process, we are presented with the opportunity to continue to further engage with them and to continue our collaboration in their program renewal efforts.
Overview of Program Evaluation Data and Plan of This Study

Data sets combined and linked

The North Carolina Department of Public Instruction (NC DPI) provided us with five large databases that contained individual student information on student characteristics, test scores, immigrants, students who have been classified as limited-English-proficient (LEP), and student participation in various types of education programs. In addition, we collected data from the participating dual language schools, with the kind cooperation of their administrators and teachers.

Using professional-level database software (FoxPro), statistical analysis software (SPSS), and spreadsheet (Excel) software, we linked individual student records from the six separate data sources, and performed the analyses found in this report. In addition, we interviewed principals and program directors in the participating dual language schools, and used this narrative data to inform and guide our quantitative analyses and to greatly improve our understanding of the phenomena we were investigating.

School districts and schools included

The data sample analyzed in this study consists of all available information on the students in the North Carolina school districts that have been operating dual language programs for at least four years. The school districts involved are: Chapel Hill-Carrboro, Charlotte-Mecklenburg, Winston-Salem-Forsyth County, Greene County, Durham County, and Chatham County. There are a total of 103,121 students in grades 3-8 in these districts who have testing records for 2008 and who were individually matched with their records from five other databases furnished by NC DPI or by the school districts themselves.

Within these districts, there are 10 schools that operate dual language programs of one of two types – either “whole-school models” in which all of the school’s students participate, or “strand” models in which at least one class per grade participates. These schools are: Carrboro Elementary and Glenwood Elementary (Chapel Hill-Carrboro), Collinswood Elementary, Oaklawn Elementary, and Smith Elem/Middle (Charlotte-Mecklenburg), Ashley Elementary (Winston-Salem/Forsyth), Snow Hill Primary and West Greene Elementary (Greene County), Southwest Elementary (Durham County), and Siler City Elementary (Chatham County).

Two schools were included in the school interviews but not in the testing analyses. Siler City’s dual language program had not reached the third grade (first year of NC state testing) as of 2008 so there were no dual language students with state test scores available. In addition, Casa Esperanza, a Raleigh charter school, was initially included but we were not able to secure the necessary school data to include this school fully in our analyses.

Organization of the Study

There are three major analytical parts of this study. First, in order to usefully describe the context in which programs for English learners operate in our participating North Carolina school districts and schools, we conducted Stage 1 analyses (see prior description) that provided information on the instructional context in participating schools. Second, we collected and analyzed descriptive data on the characteristics of the schools in the participating districts that operate dual language schools. Third, we also analyzed the achievement of these students on the North Carolina End-of-Grade state test for 2008.
Specifically, in Part I of this report, we conducted a needs-assessment focused on documenting any current achievement gaps among groups of students. Second, in Part II, we conducted an analysis of student and program characteristics by school and district, and by programs (dual language vs. non-dual language). Third, in Part III, we carried out an analysis of Reading scores, at the individual student level, from the 2008 NC state testing program for those students, schools, and districts described in Part II.

Part I--Needs Assessment

The needs assessment consists of descriptive statistical analyses conducted on the large data set which we had constructed using six data sets provided by NC DPI staff and by the participating schools. Specifically, we performed a Thomas-Collier disaggregation (see explanation of Stage 1 analyses in a prior section of this report) on the data set that had been created by linking six other data sets, in order to find out the degree of achievement gap that exists for English learners—termed limited-English-proficient (LEP) students in these findings—and other groups of interest. We sought to find out whether any achievement gaps exist among three major groups, and several sub-groups, represented by students who are currently or formerly LEP, students who are of a language-minority background but who have not been classified as LEP, and those non-language-minority students who are native-English speakers.

Context Analyses - The Thomas-Collier Disaggregation

In Stage 1 analyses, we examine the context of education in the North Carolina school districts studied by comparing the 2008 achievement levels of three major groups:

- Group 1— the English learners or limited-English-proficient (LEP) students – this group is made up of language-minority students who have been tested and classified as LEP. This group is divided into current LEPs (students currently served in programs for LEPs) and former LEPs (students who have received appropriate LEP services in the past, have been reclassified as proficient in English, and are now in the mainstream instructional program);

- Group 2—the Language-Minority-but-not-LEP students (LMNL) – these students are classified as language-minority because they have a non-English language in their background by means of ethnic group or home language. However, these students have been evaluated by the schools as not needing LEP services and thus have not been classified as LEP;

- Group 3—the Non-Language Minority students (NotLM ) – these students are not language-minority and are native-English speakers (NES). They include native-English speakers who are Caucasian, African-American, of multiple racial/ethnic backgrounds, and of all socioeconomic groups.

Since 1985, we have used this disaggregation of student test scores to determine the extent of any achievement gap between LEPs and Language-Minority-but-not-LEPs. In addition, we have compared these groups’ achievement with that of the native-English speakers to determine the size of their achievement gaps with this group. **Overall, our goal for school districts is that former LEPs and LM-but-not-LEPs should be scoring similarly to native-English speakers in terms of central tendency (average scores) as well as dispersion (‘spread’ of scores) and that all groups should be at or above grade level in achievement.** In other words, the distributions of test scores for these three groups should be similar and should generally overlap each other, thus providing evidence that the school districts’ programs have worked to reduce or eliminate achievement gaps among groups.
Of these three groups, the native-English speakers (NES) can be most diverse, depending on the local distribution of ethnic/racial groups and the range of socioeconomic status within the NES group. Initially, we calculated the average and range of performance for the overall group of native-English speakers. However, we also looked for significant sub-group differences that might lead to a possible misinterpretation of the gap closure information in the Thomas-Collier disaggregation. The current North Carolina analyses turned out to contain just such a situation.

We first ran analyses comparing the current LEPs, former LEPs, Language Minority-but-not-LEPs, and native-English speakers. The results are shown below in Figures 1 and 2.

**Figure I-1 findings:** First, Figure I-1 shows that former LEPs score higher (by about .80 standard deviations) than current LEPs, an indicator that the districts’ programs for English learners may be having a significant positive effect. However, former LEPs are not quite reaching the achievement levels of the language-minority-but-not-LEP group, indicating that the districts’ LEP programs may not be fully eliminating the achievement gap between these two groups.

Finally, the native-English speakers, labeled “Not LM” in the figure, are scoring slightly higher than former LEPs, but lower than the LM-but-not-LEP group. Since, in the researchers’ experience, it is quite atypical for language-minority students to outscore native-English speakers, we felt that this finding deserved further attention and analysis. As a first step, we performed a breakdown of the achievement scores by Grades as a more appropriate and a more valid view of the performance of the groups and sub-groups (see Figure I-2).

**Figure I-1:** 2008 Reading – All Schools, All grades By Thomas-Collier Categories
**Figure I-2 findings:** Figure I-2 breaks down student achievement by Grades across all participating schools. In this display of student achievement by Grades, the same trends are seen as in the display combining Grades (Figure I-1). Former LEPs still outscore current LEPs in all Grades as before. Also, the same trend of similar scores for former LEPs and language-minority-but-not-LEPs is seen. Finally, the unusual finding that the native-English speakers (Not LM) score slightly lower than the former LEPs and LMNLs is still present in all Grades.

**Essentially, the native-English speakers as a group are scoring lower than they should be.** The fact that the native-English-speaker group standard deviations across the grades are larger than for the other groups indicates the possible effect of another variable. We looked at several sub-groups of native-English speakers and noted that African-American students made up about 39% of the sample and Caucasian students were about 36%. Either of these two groups is large enough to substantially affect the native-English-speaker average achievement, so we further divided the NotLM group into NotLM-W (white native-English speakers), NotLM-B (black native-English speakers), and NotLM-O (native-English speakers of other racial/ethnic background), as shown in Figure I-3 below.

Figure I-3 displays the expanded categories of students for all schools and all grades combined and Figure I-4 displays the same student categories broken down by grades.
Figure I-3: 2008 Reading by Thomas-Collier Category
All Students, All Grades Combined

<table>
<thead>
<tr>
<th>LABEL</th>
<th>Group Description</th>
<th>Average Score</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Current</td>
<td>Current LEPs</td>
<td>341.0</td>
<td>10,364</td>
</tr>
<tr>
<td>Former</td>
<td>Former LEPs</td>
<td>349.0</td>
<td>3,624</td>
</tr>
<tr>
<td>LMNL</td>
<td>Language Minority but Not LEP</td>
<td>352.9</td>
<td>7,392</td>
</tr>
<tr>
<td>NotLM-B</td>
<td>Not Language Minority-Black (Native-English speakers)</td>
<td>345.7</td>
<td>40,206</td>
</tr>
<tr>
<td>NotLM-W</td>
<td>Not Language Minority-White (Native-English speakers)</td>
<td>355.7</td>
<td>37,595</td>
</tr>
<tr>
<td>NotLM-O</td>
<td>Not Language Minority-Other (Native-English speakers)</td>
<td>350.1</td>
<td>3,882</td>
</tr>
<tr>
<td>Refused</td>
<td>Refused LEP services</td>
<td>342.7</td>
<td>58</td>
</tr>
</tbody>
</table>
**Figure I-3 and I-4 Findings:** It is apparent that the scores of African-American students in the not-LM (native-English speaker) group are much lower than for Caucasian students and lower than for Not-LM students of other racial/ethnic backgrounds. In Figure I-4, the African-American students’ average scores range from 9 points to 12.3 points less than the Caucasian students’ averages across grades 3-8, a difference equivalent to about one standard deviation. In a normal distribution of scores, this difference would be equivalent to the difference between the 16th and 50th percentiles.
Across grades 3-8, the Caucasian students scored highest, the former LEPs scored next highest, and the African-American students scored lowest. Only the current LEPs and the students whose parents refused LEP services scored lower than African-American native-English speakers. Specifically, the Caucasian students outscored the former LEPs in all grades by scale score differences ranging from 2.0 - 4.9 points (equivalent to 0.2 to about 0.4 standard deviations). However, the former LEP students also outscored the African-American students in all grades, and by even larger amounts. These former LEP-African-American scale score differences ranged from 5.3 - 7.4 points (equivalent to about 0.6 - 0.7 standard deviations). Finally, the Caucasian students outscored the African-American students in all grades by scale score differences ranging from 9.0 – 12.3 points (equivalent to about 0.9 to 1.0 standard deviations).

### Part I Conclusions:

1. When compared to Caucasian native-English-speaking students, there exists a significant achievement gap for former LEPs. In other words, the school districts’ overall program offerings for LEP/English learners appear to partially close the achievement gap, but a significant gap remains.

2. Reading achievement scores of the former LEPs are approximately equal to those of the language-minority-but-not-LEPs for each Grade. Thus, the former LEP students have experienced enough gap closure to approximately match the scores of those language minority students who were deemed not to need LEP services.

3. African-American students who are not language-minority and are native-English speakers are scoring much lower in Reading than Caucasian native-English-speakers and significantly lower than the former LEPs as well. Very large achievement gaps remain for African-American native-English speakers as a group.

The needs of African-American native-English speakers are not addressed by traditional remedial English-as-a-Second-Language (ESL) and transitional bilingual education (TBE) programs that are designed and operated for LEP/English learners. In addition, language-minority-but-not-LEP students are also not usually included in remedial ESL and TBE programs. However, the gap closure needs of both language-minority-but-not-LEP students and African-American native-English speakers can be addressed in two-way dual language programs, which also focus on closing the achievement gap for LEP/English learners.
Part II -- Two-way Dual Language Schools: Overview from Interview Data

In this report, we are focusing on the ten public schools in North Carolina that have established integrated two-way dual language classes for at least the past four academic years. As researchers with 24 and 38 years of experience, respectively, in working with dual language schools, we are most impressed with the North Carolina school administrators whom we interviewed, especially in their knowledge of this program model and their commitment to authentic implementation of the most fundamental aspects of dual language programs that research has shown make a big difference in effectiveness of the program.

Important “non-negotiable components” are a minimum of 50 percent of instruction in the non-English language, strict separation of the two languages, and a K-12 commitment (Rogers, 2009). In addition, our longitudinal research (e.g. Thomas & Collier, 1997, 2002, 2009) illustrates that two-way schools reach the highest levels of achievement, with an integrated mix of students (e.g. native-Spanish-speaking and native-English-speaking students). In research studies around the U.S., this program model is dramatically successful in fully closing the achievement gap for students who begin schooling with no proficiency in English. In addition, this program has the long-term potential to substantially raise academic achievement for all other students participating in these dual language classes. After 5-6 years in the program, typically most students have reached high levels of achievement across the curriculum (grade-level and above) in both English and the non-English language.

Strand or whole school. Most of these ten NC schools have developed this program as a strand within the school, with one-half or fewer classes of each school receiving schooling through two languages, with the exception of the three schools in Charlotte-Mecklenburg County, all of which are whole-school models. Each school has grown their dual language program grade by grade, starting with kindergarten, and two schools have preschool dual language classes.

Length of dual language program. Collinswood Language Academy in Charlotte has the longest running program, completing 12 years from 1997 to 2009. The newest program among the schools included in this study is Siler City Elementary in Chatham County, with four years completed (Grades K-3) as of Summer of 2009. Eight of the ten programs have completed at least six years of Grades K-5, while Casa Esperanza of Raleigh has completed PK-4 by Summer, 2009.

Planning for middle school and high school. Nine of the ten programs have plans in place for the dual language students moving into middle school. Eight of the ten programs will have students entering a dual language middle school program next year. In Chapel Hill, the two dual language schools had their first classes of students attending sixth grade at McDougle Middle School in 2008-2009. In Charlotte, students from the Collinswood and Smith programs have reached the upper grades at West Mecklenburg High School.

Non-English languages. Most of these programs are Spanish-English programs for native-Spanish speakers and native-English speakers, while one school in Chapel Hill-Carrboro City, Glenwood Elementary, is a Mandarin Chinese-English program for native speakers of Mandarin Chinese and Cantonese as well as for native-English speakers. Oaklawn Language Academy in Charlotte has both a Spanish-English two-way program and a French-English program with some native-French speakers attending along with native-English speakers. The Oaklawn French-English program will move to Smith Language Academy in the Fall 2009 for Grades K-8. For Grades K-8, Smith Language Academy also operates classes for native-English speakers in five languages (Chinese, Japanese, German, French, or Spanish). Also for the middle school Grades 6-8, Smith receives students from the two-way Spanish-
English dual language program at Collinswood (K-5) and next year from two-way students attending Oaklawn, who just completed Grades K-5 in either the Spanish-English or French-English classes.

**Balance of English learners and native-English-speaking students.** Nine of the ten schools have close to an equal number of native speakers of the non-English language and native-English speakers at each grade level, making them balanced two-way schools serving both student populations equally. Casa Esperanza, a Raleigh public charter school, is the only school with very few native-Spanish speakers (8 percent of the student body), so it is mainly a one-way dual language program for native-English speakers.

**Special populations being served in dual language classes.** Several of these schools have enrolled significant numbers of African-American students in the dual language classes, as well as students of low socioeconomic status (as measured by free and reduced lunch), making these classes an interesting experiment in closing the academic achievement gap for these groups, as well as for the English learners attending dual language classes.

**Staff development.** All of the staff of these dual language programs have received significant amounts of training in SIOP/TWIOP strategies, and four schools have literacy coaches trained in dual language techniques who serve as ongoing in-house staff developers. Several principals described sending their teachers to summer institutes and state and national conferences to keep them up to date with the field. NC DPI-sponsored training and conferences have been very important for all of these programs for professional development and networking. Five programs maintain a close connection with the teacher education faculty at a local university in their region.

**Popularity of program.** These dual language programs are so well liked by English-speaking parents that every school has a waiting list for the English speakers hoping to attend, and eight of the ten schools enroll English-speaking students by lottery because they have such a long waiting list. For most of these schools, Hispanic and Chinese parents have become strong advocates of this program model with each additional year of maturity of the program. Three of the Spanish-English schools have had to work hard to keep up their Hispanic enrollment, because they are not located in a neighborhood with larger Hispanic demographics. Five programs are magnet district-wide programs. Most of these schools describe great support at the central administrative level, and they acknowledge that more dual language schools could be developed with the availability of qualified bilingual staff.
Characteristics of Individual Dual Language Schools: from Interview Data

Whole-school models: Charlotte-Mecklenburg Schools

Collinswood Language Academy (K-5): This Spanish-English two-way magnet school with enrollment by lottery (grouped by language spoken at home) was begun in 1997; thus it is the longest running dual language program in North Carolina. This school feeds into Smith Language Academy (6-8) and then West Mecklenburg High School (9-12), providing continuing academic work in Spanish literature and language arts for dual language students at secondary level (but not other subjects taught through Spanish). This K-12 commitment of the school district is an important principle of quality dual language programs (Rogers, 2009).

Ethnicity, exceptionality, English proficiency, and socioeconomic status (SES) of students. For the school year 2007-2008, the academic year for which this research report is examining students’ test scores, 59.4 percent of the students attending Collinswood were Hispanic, 18.8 percent Caucasian, 12.1 percent African-American, 8.3 percent multi-racial, and 1.1 percent Asian-American. English learners were 49.9 percent of the student body, 10.2 percent exceptional children, and 10.5 percent identified as academically and intellectually gifted. As a measure of SES, 54.8 percent of students were granted free or reduced-price lunch.

Percentage of instructional time in each language and integration of students. Students of both language backgrounds are taught together, with both groups receiving initial literacy development in Spanish at kindergarten level for 90 percent of the instructional time, as well as 45 minutes of English literacy development per day. Beginning in first grade and continuing through fifth grade, half of each instructional day is taught in Spanish and the other half in English. Team teaching, with one teacher representing one language, ensures that all classrooms follow the dual language principle of strict separation of the two languages (Rogers, 2009). Collinswood teaches reading, writing, and science in English while math, social studies, and Spanish literacy are taught in Spanish, with some variations as driven by books available in each language.

Teachers and staff development. Teachers at Collinswood are very multi-national and the curriculum multicultural. Teachers are supported with ongoing, differentiated professional development, with a literacy coach who serves as director of staff development.

Special characteristics. Collinswood Language Academy has been designated an Honor School of Distinction by the state of North Carolina, as well as a high quality international Spanish academy by Spain’s Ministry of Education. The school has also been evaluated by Cambridge University and the Bill Gates Foundation and received outstanding evaluations. The school is one of three south feeder area schools in the Charlotte-Mecklenburg school district with the highest enrollment of low-SES students; yet it is one of the highest achieving schools in the state.
Oaklawn Language Academy (K-5): This Spanish-English two-way magnet with enrollment by lottery was begun in 2004, to serve the north metro region and because the Collinswood waiting list was too long. The Oaklawn French-English one-way program mainly for English speakers, but with some French speakers enrolled, which co-existed with the Spanish-English Oaklawn program for the past five years, will be housed in Smith Language Academy (K-8) beginning in 2009. Students from Oaklawn and Collinswood feed into Smith for grades 6-8, for a continuation of the Spanish-English program at a 25:75 level (Spanish 25 percent; English 75 percent) and then attend West Mecklenburg High School (9-12), providing continuous K-12 commitment to high quality language study.

Ethnicity, exceptionality, English proficiency, and SES of students. For the school year 2007-2008, 57.5 percent of the students attending Oaklawn were African-American, 27.9 percent Hispanic, 7.7 percent multi-racial, 6.4 percent Caucasian, and 0.6 percent Asian-American. English learners were 24.3 percent of the student body, 4.7 percent were exceptional children, and 63.8 percent received free or reduced-price lunch.

Percentage of instructional time in each language and integration of students. In the Spanish-English program, students of both language backgrounds are taught together, with both groups receiving initial literacy development in Spanish at kindergarten and first grade levels. Two years ago the program increased the percentage of Spanish in the lower grades to 90 percent in kindergarten and 75 percent in first grade:

<table>
<thead>
<tr>
<th>Grade</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>K</td>
<td>90:10</td>
</tr>
<tr>
<td>1</td>
<td>75:25</td>
</tr>
<tr>
<td>Grades 2-5</td>
<td>50:50</td>
</tr>
</tbody>
</table>

In 2009, they will increase instructional time in Spanish in kindergarten even more, with only the specials (art, music, and physical education) being taught in English. Team teaching, with one teacher representing one language, ensures separation of the two languages for instruction. Teachers make collective decisions as an integrated learning community regarding what aspects of the curriculum will be taught in Spanish and in English, varying from year to year.

Teachers and staff development. Teachers at Oaklawn are very multi-national, their countries of origin representing North, Central, and South America, the Caribbean, and Europe. The curriculum is very cross-cultural, representing many regional variations. The staff enjoy the challenge of “finding common ground” for language and culture issues. Teachers are supported with many types of professional development, and an in-house literacy facilitator.

Special characteristics and challenges. Oaklawn Language Academy was the only World Languages program nationally to be selected for the “School of Distinction” award from the Magnet Schools of America in 2008. This school serves a large population of African-American students (57.5 percent of the student body), and 63.8 percent of their students are of lower SES background. Since Hispanics do not represent a major demographic portion of the neighborhood, and thus Spanish-speaking families have to commit to long bus rides for their children to attend this magnet school, enrollment of Spanish speakers is a challenge for the school, to maintain equal numbers of students of each language background. Furthermore, mobility of students is an issue.
Smith Language Academy (K-8): The Spanish-English two-way magnet for Grades 6-8 at Smith continues the dual language program for the students coming from Collinswood and Oaklawn (K-5). Students in this middle school program experience 25 percent of their instructional time in Spanish and 75 percent in English. The Spanish instructional time is mostly focused on language arts and literature, because they do not have the certified staff to teach other subjects in each non-English language, since this magnet school houses five instructional languages in addition to English: Japanese, Chinese, German, French (K-8), and Spanish (6-8). Students graduating from Smith attend West Mecklenburg Academy of International Languages (9-12), continuing high academic levels of language and literature in their instructional languages of Grades K-8, as well as a third language if desired.

Ethnicity, exceptionality, English proficiency, and SES of students. For the school year 2007-2008, 45 percent of the students attending Smith Language Academy were Caucasian, 27 percent African-American, 15.5 percent Hispanic, 8.8 percent multi-racial, and 3.5 percent Asian-American. English learners were 20.5 percent of the student body, 7.9 percent were exceptional students, and 30.9 percent received free or reduced-price lunch.

Percentage of instructional time in each language and integration of students. The one-way K-5 programs for mostly native-English speakers at Smith are taught intensively in the target languages, with initial literacy development in the non-English language. Except for specials in English, the French and German classes are taught only in those languages for Grades K-2. For the Japanese and Chinese K-2 classes, because of the different writing system, English literacy is taught one hour daily. Students take home books in English to develop English literacy at home. For all four languages, Grades 3-5 continue with 90 percent of the instructional day in the non-English language and only 10 percent in English, except that some of the academic texts are read in English if similar material is not available in the non-English language. At the middle school level, the core curriculum is taught in English, with students attending only language arts and literature classes in the non-English language. During the English instructional time, students across all language programs are integrated together at each grade level, K-8. A few native speakers of the non-English languages are enrolled in each language program, but we are classifying the Chinese, Japanese, German, and French programs at Smith as mainly one-way programs for native-English speakers. But at the middle school level, students from the two-way Spanish-English dual language classes at Collinswood and Oaklawn bring continuation of the two-way, integrated model with both Hispanic and English-speaking students attending the Spanish middle school classes together. Also Oaklawn’s French-English program has enough native French speakers to enrich the middle school classes as a somewhat integrated model, but not enough to be classified as two-way.

Teachers and staff development. Each teacher is a native speaker of the instructional language, and no translation is allowed. The experienced teachers mentor new teachers, and they all receive lots of support through ongoing staff development. Cross-cultural, multi-national experiences abound in this school, as in Collinswood and Oaklawn.
Dual Language strands within schools: Chapel Hill-Carrboro City Schools

Carrboro Elementary School (PK-5): This Spanish-English two-way dual language program is housed in a neighborhood school, not a magnet. The program was started in 2002, and half of the 28 classes at the school are dual language Spanish-English classes, including 3 preschool classes. The program is so popular with English-speaking families in the attendance zone that they have had to institute enrollment by lottery for the English speakers. They are able to enroll all Spanish-speaking students in the attendance zone and keep the classes half English-speaking, half Spanish-speaking. Carrboro Elementary feeds into McDougle Middle School, which has a carefully planned continuation of the dual language program, with social studies and language arts taught in Spanish. Plans are being made for continuation Spanish courses at the feeder high school. Students in the first class reached the sixth grade in 2008-2009.

Ethnicity. In 2008-09, students attending Carrboro Elementary School were 55 percent Caucasian, 22 percent Hispanic, 13 percent African-American, 7 percent multi-racial, and 3 percent Asian-American, according to the principal. In 2007-08, the percentages were 58.6 percent Caucasian, 17 percent Hispanic, 11 percent African-American, 7 percent Asian-American, and 5.6 percent multi-racial.

Percentage of instructional time in each language and integration of students. This school chose the 50:50 model, with half of the instructional time in Spanish and half in English. But in 2008-09 they created an interesting feature, with Spanish literacy taught first for Grades K-2, similar to 90:10 models, a decision that according to research strengthens literacy development in both languages in the long run. In Grade 2 some English writing is introduced. For Grades K-2, math and science are taught in English, while social studies and language arts are taught in Spanish. For Grades 3-5, every quarter the instructional language changes for these four subjects, so that students receive equal instructional time in both Spanish and English for every subject. Even the specials have been included in this formula, so that during the Spanish specials, 90 minutes per week, cultural exploration through dance, drama, music, and art of the Hispanic world takes place. In the dual language classes, Spanish-speaking and English-speaking students are integrated together at all times. Two teachers partner at each grade level, each representing one of the two instructional languages.

Teachers and staff development. Teachers at Carrboro Elementary are very multi-national. Many of the staff have attended Dual University training in the Midwest, and the experienced teachers pair with new teachers recruited from other countries. Teachers receive ongoing professional development, provided by literacy and other resource staff.

Special characteristics and challenges. For native-Spanish-speaking families, the dual language program is “fantastic” for the educational strength it provides for their children as well as the social and personal support for families. The parent groups that support the program, Padres Unidos and others, are very active in helping each other, addressing topics of great need in the community in their meetings. A challenge for the dual language program is keeping highly qualified Spanish-speaking staff, since a number of the teachers have been recruited through special visa arrangements, and they have to return to their countries of origin after several years here.
Glenwood Elementary School, Chapel Hill (K-5): This Mandarin Chinese-English two-way dual language program is located in a neighborhood school, while the Chinese dual language classes are a magnet program with district-wide enrollment by lottery. The program was started in 2002, with one class per grade level, representing one-fourth of the total school. Mandarin Chinese native speakers and Cantonese native speakers are both allowed to enroll, because although the oral languages are different, they are united by the same writing system. Glenwood Elementary feeds into McDougle Middle School, and the dual language class continues with a Chinese teacher for one period each day. Students in the first class reached the sixth grade in 2008-2009.

Ethnicity. In 2007-08, students attending Glenwood Elementary School were 42.8 percent Caucasian, 32.8 percent Asian-American, 12.6 percent African-American, 8 percent multi-racial, and 3 percent Hispanic.

Percentage of instructional time in each language and integration of students. This school chose the 50:50 model, with approximately half of the instructional time in Mandarin Chinese and half in English. The specials in English add slightly more instructional time in English than in Chinese, and all students at the school in all grades receive 30 minutes of French instruction each day. All students in the dual language classes, both native Chinese speakers and native English speakers, work together at all times, with reading and writing taught in each language from the beginning in kindergarten, and each language kept separate by instructional time. Two teachers, one English-speaking and one Chinese-speaking, partner for two grade levels. To give each language equal time when students are most alert, for one grade level Chinese is the language of the morning and English the language of the afternoon. Then for the next grade level, English is the morning language, and Chinese the afternoon language of instruction.

Teachers and staff development. The Chinese teachers have experienced big cultural shifts in adjusting to U.S. schools’ ways of doing things. Nevertheless, they have worked diligently on preparing materials each year for each grade level in Mandarin Chinese to teach the North Carolina Standard Course of Study. In addition to the dual language teachers attending state and national conferences on dual language and receiving SIOP/TWIOP training, as well as literacy training, the school has an experienced bilingual educator who serves as staff developer for the program.

Special characteristics and challenges. The biggest challenge for this program is that they are among the first in the state providing instruction in Mandarin Chinese (along with Smith Language Academy in Charlotte, which is mainly a one-way program for native-English speakers), and therefore resource support is a big issue. In comparison to Spanish, which has many instructional textbooks to choose from, there are fewer resources for published materials in Chinese that are designed for U.S. schools. Parents have been enthusiastic and very supportive of the program; however, even with a strong parent advocacy group, not all Chinese-speaking parents in the school district have chosen this program model for their children. The program is very popular with English-speaking parents and thus there is a waiting list with enrollment by lottery.
Dual Language strands within schools: Winston-Salem/Forsyth County

Ashley Elementary School (K-5): This Spanish-English two-way dual language program is located in an inner city neighborhood in a magnet school with an International Baccalaureate (IB) Primary Years Program. The dual language classes were started in 2002, with by 2008-09 two classes per grade level in the lower grades and 1 class per grade level in the upper grades, representing one-third of the total school, and continuing to grow. Approximately half of each class is native-Spanish-speaking and half native-English-speaking students. Ashley feeds into Paisley Middle School (6-8), also an IB school, and plans are being discussed regarding continuation of dual language experiences at secondary level. The first class of the dual language program attended sixth grade in 2008-09, but the number of students was not sufficient to provide a special dual language class for this year’s sixth grade students.

Ethnicity. In 2007-08, students attending Ashley Elementary School were 77.6 percent African-American, 14.9 percent Hispanic, 4.2 percent Caucasian, 2.8 percent multi-racial, and 0.5 percent Asian-American.

Percentage of instructional time in each language and integration of students. This school chose to emphasize Spanish in kindergarten, and to begin with literacy development in Spanish for all students, with an 80:20 model. In kindergarten, all students in dual language receive 4 hours of instruction through Spanish, and one hour through English, which includes specials. First grade is 70:30, and English literacy is introduced during the 30 percent English instructional time. Science, social studies, Spanish language arts, and math are taught in Spanish, and 90 minutes in English include specials, guided reading, and math. Second grade is 60:40, and Grades 3-5 are 50:50. Students work together in an integrated program, with the languages separated by the teacher signaling a switch to the other language.

Teachers and staff development. Teachers at Ashley are very multi-national, with origins from North, Central, and South America, the Caribbean, and Europe, bringing an international, cross-cultural perspective to the dual language program. Classes are currently self-contained, with each teacher providing instruction through both instructional languages, but when the program has grown sufficiently, they plan to do team teaching. Teachers have experienced lots of dual language training through state institutes and national conferences as well as training provided by an in-house trainer/coach.

Special characteristics and challenges. Ashley Elementary School was named “North Carolina School of the Year” by the Visiting International Faculty Program in April, 2007. This is the largest international exchange program linking U.S. schools and teachers worldwide. The dual language program has received media attention and is supported by the community and central administrative staff. As an inner city school with a large percentage of students receiving free or reduced lunch, this school is thriving. It is important to examine the achievement of African-American students attending the dual language classes, as well as the Hispanic students, to see the effect of the program on closing the academic achievement gap for both of these groups.
Dual Language strands within schools: Durham Public Schools

Southwest Elementary School (PK-5): This Spanish-English two-way dual language program is also located in a diverse, urban school with a large percentage of African-American students (65 percent) and 60 percent of the school’s population receiving free or reduced lunch. The dual language program is popular and thus enrollment of English speakers is by lottery, following the demographic profile of the school. However, because of mobility, students often do not stay enrolled in the program into the upper grades. The program is one-fifth of the total school, with one class for each grade level. The teachers would like to expand the program, but to have an equal number of Spanish-speaking students, they would need to recruit from outside their neighborhood. The dual language classes were started in 2002, and they have reached fifth grade level, but the number of students remaining in the program is small and there is no plan for a continuation at the middle school level.

Ethnicity and SES. In 2008-09, students attending Southwest Elementary School were 65 percent African-American, and 12 percent Hispanic, with 60 percent receiving free or reduced lunch, according to the principal. In 2007-08, the percentages were 61.3 percent African-American, 15.9 percent Caucasian, 11.5 percent Hispanic, 6 percent Asian-American, and 5 percent multi-racial.

Percentage of instructional time in each language and integration of students. This school chose the 50:50 model, with equal instructional time in each language, but initial literacy is taught in English, so there is less emphasis on the non-English language in the early grades, in contrast to the 90:10 models. For Grades K-1, science and math are taught in Spanish with a little Spanish literacy development, and language arts and social studies are taught in English. Beginning in Grade 2, a literacy block provides equal instructional time for developing reading and writing in the two languages. For Grades 3-5, the whole curriculum is taught using thematic units, taught in each language. Students work together in an integrated team-taught program, with one teacher representing each language, the model for the program for 2002-2009.

Teachers and staff development. In the past years of the program, teachers have been team teaching, sharing two grade levels, but for 2009-2010, they plan to use one teacher per class at each grade level, with that teacher providing instruction through both languages, since they have experienced a significant amount of teacher attrition, partially due to the heavy workload when working with two classes of students across two grade levels. Teachers received SIOP/TWIOP training and the school has a high percentage of teachers certified to teach academically gifted students, including strategies for differentiation. ESL staff also have worked with coaching and staff development.

Special characteristics and challenges. Both student and teacher attrition appear to be significant challenges in the dual language classes at Southwest. The school serves a diverse neighborhood with both high-income and low-income students, and there is significant student turnover among families of lower-income. Nevertheless, the program is viewed very positively within the school community.
Dual Language strands within schools: Greene County Schools

Snow Hill Primary (K-2) and West Greene Elementary (3-5): This Spanish-English dual language program is unique among the NC schools to date, since it is located in a rural context. Begun in 2003, as the program has gained in popularity, they have had to institute enrollment by lottery. In 2008-09, there were two dual language classes at each grade level out of a total of 10-13 classes for every grade, K-3, and one dual language class for each of Grades 4 and 5. For 2009-2010, plans are in place to continue the program at the middle school (6-8), with the science curriculum to be taught by a Spanish-speaking teacher who was Teacher of the Year. The plan for high school is to offer high levels of courses in Spanish, and students can take a third language.

Ethnicity and SES: In this very rural community, Hispanics were 22 percent of the Greene County school population when the program started in 2003, and they are now 25 percent of the total. For the dual language classes, enrollment is balanced as much as possible to follow the demographic profile of the total school; thus half of each class is native-Spanish speakers, half native-English speakers, with among the native-English speakers 25 percent African-American and 25 percent Caucasian. Both schools qualify as Title I schools, with around 75 percent receiving free or reduced lunch. This is a rural county with extensive poverty.

Percentage of instructional time in each language and integration of students. These schools chose the 50:50 model, with equal instructional time in each language, and literacy taught through both languages of instruction from kindergarten on. For Grades K-2, the alternate day model was chosen, with Spanish the language of instruction for one full day, followed by English as the language of instruction for the next full day, and alternating back to Spanish, etc. By fourth grade, alternation shifts to half-days, conducted in units. So, for example, 4 weeks of science would be taught in one language, followed by 4 weeks of science in the other language. The two languages are separated by team teaching partnerships. For Grades K-3, each teacher has one grade level and two classes; whereas teachers at Grades 4-5 have two grade levels and two classes. This will change to a more manageable teaching assignment as more classes are added, because the program continues to grow. Native-Spanish-speaking and native-English-speaking students are taught together at all times.

Teachers and staff development. Teachers in the program have received much training from a close relationship with East Carolina University faculty, as well as several years of SIOP training by an in-house trainer/coach. Several dual language teachers attended Dual University in the Midwest.

Special characteristics and challenges. In 2006, Greene County was one of four districts in the state to be designated a Global Communicators School District, a project of Governor Easley’s North Carolina in the World initiative. The county has developed these bilingual/bicultural innovations in a rural, high poverty region of the state, serving Hispanic, African-American, and Caucasian students equally. The high turnover rate of teachers in this rural community is the greatest challenge. With budget cuts during the current economic downturn, the county lost several good teachers in the dual language program, and finding qualified teachers to teach the curriculum through Spanish is challenging. Another challenge in the initial years of the program was convincing Spanish-speaking parents to enroll their children in dual language, but in recent years, the parent advocacy groups are growing. Central administration is very supportive of the dual language program.
Dual Language strands within schools: Chatham County Schools

Siler City Elementary School (K-4): This Title I school began their dual language program in 2005, thus completing Grades K-3 by 2008-09. The program is growing each year, with two kindergarten and two first grade classes, and one each for Grades 2-3 as of 2008-09. When the dual language program started in this neighborhood school, native Spanish speakers were 75 percent of the school population, reduced to 62 percent after a new school was built. To enroll a balance of native-English speakers, the program started with district-wide enrollment for this group, but as the program has gained in popularity, there is now a waiting list for native-English speakers. The feeder middle school is planning a continuation of the dual language program.

Ethnicity and SES. In 2008-09, students attending Siler City Elementary School were 62 percent Hispanic, and 83 percent of the students in the school were receiving free or reduced lunch, according to the principal.

Percentage of instructional time in each language and integration of students. This school chose to do a 90:10 balance of Spanish-English instruction for kindergarten, followed by 50:50 for the remaining grades. Thus Spanish is emphasized in kindergarten for 90 percent of the instructional time, including Spanish literacy. English literacy is begun in kindergarten for approximately 30 minutes per day, plus the specials, but is emphasized more beginning in first grade. Team teaching is done for the grades that have two dual language classes, and the two instructional languages are kept separate by alternating the curricular schedule. Siler City Elementary dual language staff confer with the Greene County dual language staff, with similar demographics and experience, and they have both adopted the terms “Spanish world” and “English world” to clearly separate the two languages of instruction. Native-Spanish speakers and native-English speakers work together.

Teachers and staff development. Teachers at Siler City Elementary have a working relationship with UNC-Greensboro, as a school site for student teachers, and for ongoing staff development. Dual language teachers have been trained in SIOP and have attended state-sponsored dual language conferences. The school has hired several Visiting International Faculty for their Spanish-speaking teachers.

Special characteristics and challenges. Community support varies, but Spanish-speaking parents in the neighborhood are becoming an advocacy group, recognizing the importance of their children developing biliteracy. With student achievement increasing each school year, the program is helping to bridge the cultural divide in the community. In this Title I school, in 2008-09 they had 11 students, including 5 Spanish speakers, qualify for academically gifted programs. Given that 83 percent of the students are of low-SES background, it will be important to follow the Hispanic students across the school years, to see the effect of the program on closing their achievement gap with native-English speakers.
Dual Language strands within schools: Raleigh, NC Public Charter School

Casa Esperanza Montessori Charter (PK-6): We are including this school in our reports of interview data, but we are not including analyses of student achievement data for this school, because we have not yet received information regarding which students have been enrolled in the dual language program. The dual language program at Casa Esperanza is what we would classify as a one-way program mainly for native-English speakers, since only eight percent of the students in the Spanish-English dual language classes are Hispanic. Casa Esperanza is a public charter school with open enrollment for students from several school districts within the Raleigh metropolitan area. Since the school districts do not provide busing to the school, parents must commit to drive their children to the school. The school is not located in a neighborhood with large numbers of Hispanic residents and thus the low Hispanic enrollment. Almost half (7 of a total of 15) of the classes at the school are Spanish-English dual language classes, and in addition all students at the school receive 30-45 minutes of Spanish enrichment daily. Enrollment is by lottery, and there is a long waiting list. As parents have requested, the school eventually plans to serve Grades PK-8, including a continuation of the dual language program.

Percentage of instructional time in each language and integration of students. While the kindergarten dual language classes are described as 90:10, reading is taught first in the students’ native tongue, so this model does not follow the classic program (first introduced in Canada in the 1960s) in which reading is taught first in the non-English language. At Casa Esperanza, native-English speakers switch to learning to read and write in Spanish when their English reading reaches a level beyond decoding, usually in first grade in this very individualized Montessori curriculum. For Grades 1-3, following a 75:25 model, social studies and science are taught in Spanish, while English language arts and math are taught in English. From Grade 4 on, instruction is 50:50 in each language, with math and science taught in Spanish, and English language arts and social studies in English. As mentioned above, only eight percent of the students are native-Spanish speakers, so this is a program mostly for native-English speakers.

Teachers and staff development. Teachers are taught to separate the languages of instruction, and teachers do team planning, with curricular materials available influencing which language is chosen for each subject. There are two native-Spanish-speaking teachers in each dual language class, and some classes also have a native-English-speaking teacher. Teachers are trained in Montessori methodology. In the other half of the school, Spanish-speaking paraprofessionals also talk and teach in Spanish, so students in all classes throughout the school get exposure to Spanish.

Special characteristics and challenges. There is strong community support for this school, and the waiting list is long. Spanish-speaking students could benefit greatly from this type of school, if busing could be provided. Since a larger number of native-Spanish-speaking teachers have been hired at this school than is typical of most of the NC dual language programs, a second school could be developed, to be located near the largest Spanish-speaking neighborhoods in the Raleigh metropolitan area, to serve the needs of Hispanic students.
Part III - Analyses from Databases By Education Programs

The third part of this study reflects analyses appropriate to Stage 3 of our overall five-stage evaluation plan that guides our long-term work with education agencies (see prior section entitled “Five Stages of Analysis”). In this third part, we examine the NC End-of-Grade test scores in Reading for school year 2007-08 (latest available data) in Grades 3-8, and break these down by programs (dual language classes compared to non-dual language classes). We also examine student sub-groups of interest within each program type.

At this stage in our overall analysis plan, these results are statistically descriptive rather than inferential and do not yet control for the possible effects of extraneous variables (as in later Stage 5 analyses). However, these test score breakdowns offer teachers, administrators, and policy makers some preliminary indicators and guidance as to the effectiveness of programs for English learners, and for other students as well, both in dual language classes and in non-dual language classes. The intent of these analyses is to provide information to decision-makers which will allow for an assessment of interim student outcomes as NC dual language programs are developing and maturing. In addition, these analyses provide decision-makers with short-term information that allows them to make informed “go vs. no-go” decisions to continue and perhaps to expand the implementation of dual language programs in North Carolina school districts.

Categories of Education Programs

The testing and other data from the combined databases allows us to distinguish three categories of education programs for English learners, native-English speakers, and other students in North Carolina. First, there are students who are receiving dual language instruction in dual language classrooms in one of the ten dual language schools found in one of the six participating school districts in this study. These students are denoted as “DL-DL” (students in a dual language school and dual language classroom) in the following graphs.

Second, for Grades 3-5, there are also students who are in the same dual language schools but not in a dual language classroom, denoted as “DL-NDL” (dual language school but not in a dual language class). These schools operate dual language strands -- one or more dual language classrooms in each Grade that are accompanied by one or more non-dual language classrooms in each Grade of that school. Thus, as an example, a school might have two dual language classrooms (DL-DL) and two non-dual language classrooms (DL-NDL) in each Grade of that school. This group represents a useful comparison group for the dual language classes in each school, since these non-dual language students are in the same school as the dual language students and thus are affected by many of the same school-wide and community-wide factors as the dual language students in that same school.

Third, there are students who are not in a dual language school at all, and thus are obviously not in a dual language classroom. These are denoted as “NDL-NDL” (non-dual language school, not in a dual language classroom). In grades 3-5, these students are in different schools than those schools operating dual language classes, but are within the same school districts as the dual language schools. In Grades 6-8, dual language programs are operated only in one school district in this study, so in these Grades, this NDL-NDL comparison group may also include students from different districts where there no dual language middle schools. In this case, these students may be affected by different school-wide and community-wide factors in grades 6-8 but have the same district-wide influences in Grades 3-5 as do the first two groups in dual language schools. This third group also provides a comparison group of sorts, but the second group (DL school-not in DL program) shares more background factors with dual language
students (DL school and DL program) than does the third group and thus the second group is of more interest when comparing achievement outcomes with the scores from dual language classes and schools.

Analyses of Testing Outcomes By Program Categories

In Figure III-1, students’ Reading test scores are displayed by Program Category (DL-DL shown in red, DL-NDL shown in green, and NDL-NDL shown in blue) as explained above, and by student Grade in school. The mean Reading achievement scores are depicted by bar graphs for each sub-group and the sub-group N’s are shown in the group labels to the right of each figure. For example, in Figure III-1, the label to the right of the red rectangle (“DL schl & DL class”) indicates that the total N for this group is 1122, and that within this group, the subgroup N’s for Grades 3,4,5,6,7 and 8 are 313,298,167,139,113, and 92 respectively.

Figure III-1 Findings:
(1) For Grades 3-5, where all three categories of students are found, the students in dual language schools and dual language classrooms outscore the other two categories by significant amounts, about 3-5 scale score points. In particular, the dual language students (DL school-DL class) outscore their comparison group, the non-dual language students (DL school, non-DL class) in the same schools. These differences correspond to approximately one-third of a standard deviation (S.D.) to about one-half of a S.D. and represent “actionable, pragmatically significant differences” in favor of the dual language classrooms.

(2) In Grades 6-8, the participating dual language middle schools are “whole-school” dual language programs, so all students in these schools attend dual language classrooms. Therefore, only two groups appear in the display for Grades 6-8; the dual language students (DL school-DL classroom) and the non-dual language students in different schools and districts (Non-DL-school, non-DL classroom). Once again, the dual language students significantly outscore the non-dual language students.

Figure III-1: 2008 Reading by Grade by Type of Class
Comparison with statewide average scores

Table III-1 below provides what Figure III-1 does not -- a comparison of the three student groups’ average achievement to 2008 statewide average scores in Reading. A comparison of the average achievement of dual language students, non-dual language students in dual language schools, and non-dual language students in non-dual language schools serves to “calibrate” the average scores depicted in Figure III-1 by providing information as to whether the three groups’ scores are below, at, or above the statewide average scores for each Grade, and by how much they differ from the statewide averages. Although the data presented is in cross-sectional form (since all data is from the 2008 testing), and is not longitudinal across Grades, a comparison of the average 2008 reading achievement of dual language students to statewide averages for each Grade is of substantial interest to those who wish to evaluate the dual language programs.

Table III-1: Average statewide and group scores in Reading for 2008

<table>
<thead>
<tr>
<th>Grade</th>
<th>Mean Scale Score</th>
<th>Standard Deviation</th>
<th>N tested</th>
<th>DL Students in DL schools</th>
<th>N</th>
<th>Non-DL students in DL schools</th>
<th>N</th>
<th>Non-DL students in non-DL schools</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>338.7</td>
<td>11.9</td>
<td>108,942</td>
<td>341.6</td>
<td>313</td>
<td>336.3</td>
<td>675</td>
<td>338.6</td>
<td>17,777</td>
</tr>
<tr>
<td>4</td>
<td>345.3</td>
<td>10.1</td>
<td>105,291</td>
<td>348.1</td>
<td>298</td>
<td>344.9</td>
<td>617</td>
<td>345.2</td>
<td>16,582</td>
</tr>
<tr>
<td>5</td>
<td>350.0</td>
<td>9.4</td>
<td>103,722</td>
<td>352.7</td>
<td>167</td>
<td>349.9</td>
<td>754</td>
<td>349.9</td>
<td>16,259</td>
</tr>
<tr>
<td>6</td>
<td>353.0</td>
<td>9.6</td>
<td>101,722</td>
<td>356.7</td>
<td>139</td>
<td>---</td>
<td>---</td>
<td>352.5</td>
<td>16,031</td>
</tr>
<tr>
<td>7</td>
<td>355.7</td>
<td>9.2</td>
<td>102,965</td>
<td>358.9</td>
<td>113</td>
<td>---</td>
<td>---</td>
<td>355.0</td>
<td>16,360</td>
</tr>
<tr>
<td>8</td>
<td>358.5</td>
<td>8.8</td>
<td>104,103</td>
<td>360.6</td>
<td>92</td>
<td>---</td>
<td>---</td>
<td>357.9</td>
<td>16,944</td>
</tr>
</tbody>
</table>

Table III-1 findings:

First, it is quite noteworthy that the DL students’ average reading scores exceed the statewide average scores in each Grade whereas the average scores of the two non-DL groups do not. Second, the DL average reading scores for each Grade exceed their corresponding statewide average scores by a range of 2.1 to 3.7 scale score points, an amount that is equivalent to .24 to .38 statewide standard deviations. These differences are large enough to not be the result of chance or measurement uncertainty, but instead to be the result of real differences in achievement between the dual language students and their non-dual language counterparts.

Third, we note that after Grade 4, the dual language students approach and then exceed the average statewide scores of students who are a year ahead of them in school. Specifically, in Grade 5, the dual language students are approaching the Grade 6 statewide average reading score. In Grade 6, the dual language students exceed the Grade 7 statewide score. In Grade 7, they also exceed the statewide average score of students who are in Grade 8, a year ahead of them in schooling.
Achievement Breakdowns by Defined Student Groups

In Figure III-1, all defined student groups, consisting of current LEPs, former LEPs, students who are language minority but not LEP, and all native-English speakers (African-American, Caucasian, and other), were combined for an overall look at the performance of students in dual language classrooms compared to students not in dual language classrooms. However, in Figures III-2 through III-7, the average reading achievement in each Grade is presented separately for each defined student group.

The mean Reading achievement scores are depicted by bar graphs for each sub-group and the sub-group N’s are shown in the group labels to the right of each figure, as in Figure III-1. These displays allow us to see how each of the defined student groups fared in the three instructional environments (dual language class in dual language school, non-dual language class in a dual language school, and non-dual languages classes/schools).

Figure III-2 Findings for current LEP students:
(1) In Figure III-2, we examine the performance of current LEP students who were tested using the End-of-Grade tests in 2008. We note that students in dual language classes outscore both non-dual language classes in the same schools and they outscore non-dual language classes in non-dual language schools in each Grade from Grade 3-8. These differences, in favor of dual language students, are significant in Grades 3, 5, 6, and 7, with a small but non-significant difference in Grade 4. In Grade 8, the number of current LEP students in dual language classes is too small for reliable analysis (n=9).

(2) Overall, the current LEP students in dual language classes are prospering academically, as measured by the Reading subtest of the North Carolina End-of-Grade tests for each Grade, when compared to current LEPs in non-dual language classrooms. At the same time, they are acquiring deep proficiency in two languages, English and their first language (e.g., Spanish, Mandarin Chinese) at no cost to their academic development while the current LEPs in non-dual language classrooms are mostly scoring lower on the EOG tests and are acquiring one language only (English).

Figure III-3 Findings for former LEP students:
(1) In Figure III-3, we examine the performance of former LEP students (LEP students who have been reclassified into the instructional mainstream) who were tested using the End-of-Grade tests in 2008. Students in dual language classes tend to score at or above the levels of non-dual language classes in the same schools, but the group N’s are small (less than 15) in Grades 5-8. Former LEPs in dual language classes also score at or above the scores of those in non-dual language classes in non-dual language schools in each Grade from Grade 3-8, but again the dual language N’s are small in Grades 5-8.

(2) Overall, the former LEP students in dual language classes are scoring at the levels of the non-dual language students in the same schools in Grades 3 and 4. In Grades 5-8, they tend to outscore non-dual language students but the group N’s are too small for reliable analysis. Former LEP students equal the performance of non-dual language students who are also former LEPs, while acquiring deep proficiency in two languages rather than just one, and at no cost to their academic development.
Figure III-2: 2008 Reading by Grade by Type of Class (Current LEPs only, N=10,364)

Figure III-3: 2008 Reading by Grade by Type of Class (Former LEPs only, N=3,624)
Figure III-4 Findings for language minority students who are not LEP:
(1) In Figure III-4, we examine the performance of students who are language minority but who have not been classified as LEP (LM-not-LEP) by the school districts, and who were tested using the End-of-Grade tests in 2008. LM-not-LEP students in dual language classes tend to score at or above the levels of non-dual language classes in the same schools except in Grade 5, a rare exception to the trend established in other Grades and in other group comparisons. However, the group N’s for Grades 6-8 are small, making reliable analyses difficult. Former LEPs in dual language classes also score at or above the scores of those in non-dual language classes in non-dual language schools in each Grade from Grade 3-8, but again the dual language N’s are small in Grades 6-8.

(2) Overall, the LM-but-not-LEP students in dual language classes are scoring at the levels of the non-dual language students in the same schools in Grades 3 and 4 and below the levels of non-dual language students in Grade 5. Since this Grade 5 performance is not found in any of the other groups’ comparisons, it may be the result of chance or of some factor applying only to the 23 students involved. In general, LM-but-not-LEP appear to be receiving benefits from dual language instruction.

Figure III-4: 2008 Reading by Grade by Type of Class
(Language Minority but not LEPs only, N=7,450)

Figure III-5 Findings for African-American native-English speakers:
In Figure III-5, we examine the performance of students who are African-American native-English speakers and who were tested using the End-of-Grade tests in 2008. Previous analyses in Part I of this report have identified this group as having significant achievement gaps when compared to Caucasian native-English speakers. Thus, it is of great interest to inquire whether African-American native-English speakers in dual language classes are achieving at levels above or below their counterparts in non-dual language classes.

(1) African-American students who are not language minority and are native-English speakers excel in dual language classrooms, especially when compared to their non-dual language counterparts in the same schools. In this case, African-American students in dual language classes score 6.5 (Grade 3),
5.0 (Grade 4), and 7.2 (Grade 5) scale score points higher than their same-school counterparts in non-DL classes, an amount equivalent to a range from .50-.77 standard deviations for these Grades. When compared to their non-DL counterparts in non-DL schools, the DL African-American students score 4.4 (Grade 3), 3.4 (Grade 4), and 5.7 (Grade 5) scale score points higher, differences that amount to a range of .34 to .60 standard deviations for those Grades. The N’s for these comparisons are all 40 or above, and so these differences may be interpreted as substantial and significant in terms of effect size.

In Grades 6-8, the dual language student N’s for African-American students cluster around 30, enough for reliable analysis. In this case, the dual language African-American students outscore their counterparts in non-dual language middle schools by 3.8 points (.40 s.d.) in Grade 6, 4.8 points (.52 s.d.) in Grade 7, and 2.5 points (.28 s.d.) in Grade 8. These differences are of practical significance and favor the dual language African-American students.

(2) Overall, it appears that African-American native-English speakers’ instructional needs are being substantially met in dual language classrooms, and met to a significantly lesser degree in non-dual language classrooms.

**Figure III-5: 2008 Reading by Grade by Type of Class**
*(Native-English Speakers-Blacks Only, N=40,206)*
Table III-2: Average statewide and Black student scores in Reading for 2008

<table>
<thead>
<tr>
<th>Grade</th>
<th>Mean Scale Score</th>
<th>Standard Deviation</th>
<th>N tested</th>
<th>DL Students in DL schools</th>
<th>N</th>
<th>Non-DL students in DL schools</th>
<th>N</th>
<th>Non-DL students in non-DL schools</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>338.7</td>
<td>11.9</td>
<td>108,942</td>
<td>338.1</td>
<td>77</td>
<td>331.6</td>
<td>309</td>
<td>333.7</td>
<td>6,584</td>
</tr>
<tr>
<td>4</td>
<td>345.3</td>
<td>10.1</td>
<td>105,291</td>
<td>344.3</td>
<td>68</td>
<td>339.3</td>
<td>242</td>
<td>340.9</td>
<td>6,214</td>
</tr>
<tr>
<td>5</td>
<td>350.0</td>
<td>9.4</td>
<td>103,722</td>
<td>351.7</td>
<td>40</td>
<td>344.5</td>
<td>232</td>
<td>346.0</td>
<td>6,318</td>
</tr>
<tr>
<td>6</td>
<td>353.0</td>
<td>9.6</td>
<td>101,722</td>
<td>351.9</td>
<td>32</td>
<td>------</td>
<td></td>
<td>348.1</td>
<td>6,275</td>
</tr>
<tr>
<td>7</td>
<td>355.7</td>
<td>9.2</td>
<td>102,965</td>
<td>356.0</td>
<td>29</td>
<td>------</td>
<td></td>
<td>351.2</td>
<td>6,684</td>
</tr>
<tr>
<td>8</td>
<td>358.5</td>
<td>8.8</td>
<td>104,103</td>
<td>356.6</td>
<td>30</td>
<td>------</td>
<td></td>
<td>354.1</td>
<td>7,072</td>
</tr>
</tbody>
</table>

Table III-2 findings:
A comparison of the score of African-American native-English speaking students in dual language classes to the statewide average scores for each Grade provides information on gap closure for African-American students in dual language classes.

(1) In Grades 3, 4, 6, and 7, the African-American DL students score within 1.1 scale score points of the statewide average for that Grade, and score 1.7 points higher than the state average in Grade 5. In Grade 8, they score 1.9 points lower than the state average.

(2) However, all of these differences, whether positive or negative, are too small to be declared significant, so given the sample sizes involved, a defensible finding is that there are no significant differences (in terms of effect size) between the African-American DL students’ average reading scores and the statewide average scores for each Grade.

In other words, in general, the African-American native-English speaking students who participate in dual language classes are at or near Grade level achievement, as defined by the statewide average scale scores for each Grade. On the other hand, the African-American native-English speaking students who are in non-DL classes, whether in the same schools or in other non-DL schools, are scoring substantially below grade level, by amounts equivalent to more than a year’s difference, in terms of the grade-by-grade increases in average statewide scores.

Taken collectively, these score patterns suggest that participation in dual language programs by African-American native-English speaking students in North Carolina schools is associated with substantially higher test scores than is non-participation. If confirmed by other analyses, the implication is that promoting and developing dual language programs is a powerful way of addressing the large achievement gap experienced by African-American native-English speaking students in general. It should be noted that this interpretation is supported and confirmed by other large-scale, longitudinal research conducted in large inner-city school districts with substantial African-American student population, such as Houston, Texas (Thomas & Collier, 2002).

Figure III-6 Findings for Caucasian native-English speakers:
In Figure III-6, we examine the performance of students who are Caucasian native-English speakers and who were tested using the End-of-Grade tests in 2008. Historically, Caucasian native-English speakers are the highest scoring group when school district test scores are disaggregated. Thus, they
provide a “reference group” to which other groups are compared when gap closure analyses are performed.

A two-way dual language program provides instruction in two languages not only to language minority students (LEPs and LM-not-LEPs) but also to members of the language majority (native-English speakers in the North Carolina context). Thus, not only are we interested in how African-American native-English speakers achieve in DL programs, but also we want to know if Caucasian native-English speakers who participate will outscore those comparable Caucasian native-English speakers who do not receive dual language instruction. In addition, parents of native-English speaking children who are not familiar with the workings of dual language instruction are often concerned that their children may somehow suffer academically while mastering the curriculum in two languages rather than one.

(1) Figure III-6 shows that Caucasian native-English speakers who are not language minority (Not LM-Caucasian) and who participate in dual language classes significantly outscore their same-school Caucasian counterparts in Grades 3 and 5 by 5.3 and 3.2 scale score points respectively. In Grade 4, the same-school score difference also favors the dual language students, but is not of practical significance.

In comparison to the non-dual language students in different schools, the dual language Caucasian students also score significantly higher in Grades 3, 5, and 6. In Grades 4, 7, and 8, the dual language Caucasians score numerically higher, but not significantly so.

(2) Overall, it appears that Caucasian native-English speakers’ instructional needs are being substantially met in dual language classrooms. The Caucasian students in DL classes score numerically higher than the Caucasian students in non-DL classes in all comparisons, and there is a tendency (but not a clear pattern) toward significant differences in favor of the dual language Caucasian students (five out of nine group comparisons).

Caucasian students in dual language classes are certainly not being hurt by their participation in the dual language program (as measured by the EOG state test scores), and there is some evidence that their scores are slightly elevated in the DL classrooms. In any case, all of these Caucasian students are acquiring deep proficiency in their first language English, and in a second language, and that is a large benefit (compared to their non-DL counterparts) that comes at no apparent cost to their classroom achievement and performance in English.
Table III-3 Findings:

As in the case of other defined student groups, it is worthwhile to compare DL Caucasian students’ scores to the state average scores for each Grade.

1) Caucasian students in dual language classes score higher than the statewide average for each Grade by amounts ranging from two-thirds of a statewide standard deviation to more than 1 S.D. (Grade 5). Caucasian students in non-dual language classes also score higher than the state averages for each Grade, but to a lesser degree than the DL students.

2) Overall, Caucasian native-English speakers in all three groups examined score higher than the state averages for each Grade, and the dual language group scores highest of all. There is no evidence that participation in dual language programs hurts the achievement of these students, as compared to statewide averages, and they do acquire deep proficiency in a second language as well. In addition, there is some evidence that small-to-moderate achievement gains accompany DL program participation, as compared to those students not in DL programs.

In other large-scale research (Thomas & Collier, 2002), the authors have found that native-English speakers who participate in DL programs through the elementary school years can outgain comparable native-English speakers not in DL programs by about 7-8 national percentiles on a norm-referenced test, when tested at the end of Grade 5. The cross-sectional findings here suggest the presence of a similar effect for North Carolina native-English speakers, a phenomenon which should be followed over time as future years of test data become available.
### Table III-3: Average statewide and White student scores in Reading for 2008

<table>
<thead>
<tr>
<th>Grade</th>
<th>Mean Scale Score</th>
<th>Standard Deviation</th>
<th>N tested</th>
<th>DL Students in DL schools</th>
<th>N</th>
<th>Non-DL students in DL schools</th>
<th>N</th>
<th>Non-DL students in non-DL schools</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>338.7</td>
<td>11.9</td>
<td>108,942</td>
<td>349.2</td>
<td>80</td>
<td>343.9</td>
<td>188</td>
<td>345.9</td>
<td>6,389</td>
</tr>
<tr>
<td>4</td>
<td>345.3</td>
<td>10.1</td>
<td>105,291</td>
<td>352.2</td>
<td>82</td>
<td>351.9</td>
<td>198</td>
<td>351.3</td>
<td>6,084</td>
</tr>
<tr>
<td>5</td>
<td>350.0</td>
<td>9.4</td>
<td>103,722</td>
<td>359.6</td>
<td>41</td>
<td>356.4</td>
<td>281</td>
<td>355.4</td>
<td>6,016</td>
</tr>
<tr>
<td>6</td>
<td>353.0</td>
<td>9.6</td>
<td>101,722</td>
<td>361.7</td>
<td>49</td>
<td>-----</td>
<td></td>
<td>358.4</td>
<td>5,917</td>
</tr>
<tr>
<td>7</td>
<td>355.7</td>
<td>9.2</td>
<td>102,965</td>
<td>362.5</td>
<td>41</td>
<td>-----</td>
<td></td>
<td>360.3</td>
<td>5,892</td>
</tr>
<tr>
<td>8</td>
<td>358.5</td>
<td>8.8</td>
<td>104,103</td>
<td>364.4</td>
<td>35</td>
<td>-----</td>
<td></td>
<td>363.4</td>
<td>6,302</td>
</tr>
</tbody>
</table>

**Figure III-7 Findings for native-English speakers of Other racial/ethnic background:**

Students in this category are **non-language minority students (native-English speakers) who are reported as not African-American and not Caucasian**. This group includes students of multiple racial/ethnic backgrounds. As in the case of previous groups, it is of interest to investigate how this sub-group performs in one of three classroom conditions – in dual language classrooms operated in dual language schools, in non-dual language classrooms within dual language schools, and in non-dual language schools.

(1) However, this is the smallest of the defined student groups examined, and the N’s for the dual language students in Grades 3-8 are 18 at highest and 2 at lowest. Thus, these dual language groups are mostly too small for consistent analysis and for reliable inter-group comparison. Figure III-7 presents the numerical data for each of the three groups, with the group N’s (and Grade sub-group N’s) included within the text of the group labels on the right side of the figure.

(2) The non-dual language native-English speakers of Other background who attend dual language schools (DL-NDL) mostly score above the statewide average scores, reflecting at-or-above grade-level achievement. Those non-DL native-English speaking Others who are in non-DL schools (NDL-NDL) score slightly lower than those in DL schools in Grades 3-5 but are very near to the statewide average scores in each Grade.
Figure III-7: 2008 Reading by Grade by Type of Class
(Native-English Speakers-Others only, N=3,882)
Summary of Findings From Parts I, II, and III

Part I—Needs Assessment and Student Performance by Groups (All Programs Combined)--
Findings and Discussion:

(1) In each Grade, **Reading achievement scores of the former LEPs** (graduates of school district programs for English learners) are approximately equal to those of the students who are language minority but not classified as LEP (LM-but-not-LEP), after receiving several years of services from programs for English learners (e.g., ESL). Thus, the former LEP students have experienced enough gap closure to approximately match the scores of those language minority students who were deemed not to need LEP services.

The fact that this matching does occur is an indicator that the programs received by English learners are at least partially closing their achievement gap, by raising the scores of English learners (LEPs) to the levels of other language minority students. However, this does not address the English learners’ achievement gap with native-English speakers.

(2) In each Grade except Grade 7, **former LEPs show a significant remaining achievement gap when compared to Caucasian native-English speakers**. In other words, the school districts’ programs for English learners appear to partially close the achievement gap but a significant gap remains. Thus, these former LEPs are entering the instructional mainstream at an achievement level that does not match that of the native-English speakers. In other research studies, the typical result is that these former LEP students fail to make the same one-year’s-progress-in-one-year’s-time that is characteristic of the native-English speakers, and thus the gap slowly widens with each passing year after the LEPs re-enter the instructional mainstream.

(3) It should be noted that African-American native-English speakers (who are not language minority) are scoring much lower than Caucasians across the Grades. In addition, the African-American native-English speakers score even lower on average than the former LEPs (by about one-half of a standard deviation, or more). **It is apparent that very large achievement gaps remain for African-American native-English speakers when compared to Caucasian native-English speakers, as well as when compared to former LEPs.**

(4) The “Refused” category of students consists of language minority students whom the school district diagnosed as needing English learner services but whose parents refused those services in favor of placing their child in the instructional mainstream with a mainstream teacher who typically is not specially trained to work with English learners. The total number of “Refused” students who took the End-of-Grade tests (others took alternative assessments) is only 58, a number too small to reliably analyze by Grade. However, when combined across Grades 3-8 (N=58), **it should be noted that these students scored near the levels of the current LEPs, not near the levels of former LEPs, and certainly not near the levels of Caucasian native-English speakers.** This is an indicator that the decision to place these students in the mainstream classroom was almost certainly unwise, and that this decision is likely to result in a continued pattern of low achievement during their school years, as one cannot reasonably expect these students consistently to make one-year’s-progress-in-one-year’s-time, the standard necessary for gap closure across the Grades.

**Fortunately, the number of these students is quite small,** when compared to the numbers of current and former LEPs who did receive the “meaningful education” services that *Lau v. Nichols* (U.S. Supreme Court, 1974) specifies for English learners. This indicates that the school districts have been largely
successful in convincing parents of the great importance of allowing children who need these services to receive them, and the districts should be congratulated for that public service.

**Part II—School and Program Implementation Issues—Findings and Discussion:**

(1) As researchers with 24 and 38 years of experience, respectively, in working with dual language schools, we are most impressed with the North Carolina school administrators whom we interviewed, in their knowledge of this program model and their commitment to authentic implementation of the most fundamental aspects of dual language programs that research has shown make a big difference in effectiveness of the program. These features include:

(a) **A minimum of 50 percent of the curricular instruction taught through the non-English language, preferably up to 90 percent in the early Grades:** Research shows that 90:10 models are the most efficient in reaching grade-level achievement in English, with reading taught first through the non-English language, 90 percent of the instructional day in kindergarten, reaching a 50:50 balance of the two instructional languages by Grade 4. Seven of the ten schools emphasize Spanish in a 90:10 model for kindergarten, recognizing the importance of developing a strong reading foundation in the language less supported by the society, followed by a 50:50 model by Grade 1 or 2 or 3. The other three schools are 50:50 models. Consistent with other research studies on dual language, all students’ achievement in English in the NC dual language classes is very high, supporting this administrative decision, even given less instructional time in English. Students are receiving the mainstream curriculum taught through two languages, and they are out-scoring their peers taught only in English.

(b) **Balance of English learners and native-English-speaking students:** Nine of the ten schools have close to an equal number of native speakers of the non-English language and native-English speakers, making them balanced two-way schools serving both student populations equally. This requires a commitment on the administrators’ part to serve all student groups fairly. Several administrators of these programs also work to balance the percentages of students by each sub-group within the demographics of their schools, to give those of each ethnic group and each socioeconomic class an equal chance to participate in the program. In addition to issues of equal educational opportunity, integrating the Spanish speakers and English speakers assists with the process of acquiring the curriculum through the two languages, as each language group serves as peer tutors for the other group when their native language is used.

(c) **Integration of the two language groups:** In all ten schools, both language groups work together at all times. This illustrates NC administrators’ strong commitment to the two-way dual language model. Longitudinal research finds that all students attending this integrated form of dual language instruction reach the highest levels of English achievement across the curriculum of any program model developed for LEP/English learners. Eventually, achievement for all students in two-way dual language programs is similar to that of a gifted program.

(d) **Separation of the two languages:** An important component of high quality dual language programs is strict separation of the two languages. Research on transitional bilingual education (a program model less effective than dual language) found that bilingual teachers using lots of translation and switching back and forth between the two languages led to less development of deep proficiency in both instructional languages. Each NC dual language
program has worked hard to implement curricular separation of the two languages, mostly through team teaching, with (for example) one native-Spanish-speaking teacher and one native-English-speaking teacher sharing two classes at one Grade level. Two schools’ programs are still small and these teachers have to teach two Grade levels as well as two classes, which is a challenging teaching assignment. In each school, when the switch is made to the other instructional language, students are given a clear designation, such as moving to a different classroom (Spanish world, English world) or an instructional signal given in the other language, or simply by the teacher, who speaks only the one language with students.

(e) **Continuation of the dual language program into the middle school years:** Nine of the ten programs have plans in place for the dual language students moving into middle school, and in the Fall of 2009, eight of the ten programs will have dual language students entering a dual language middle school program. Research has found that when the program is continued throughout the elementary school years and into the middle school, all student groups will reach grade-level and frequently above-grade-level achievement in English, through the stimulus of schooling through two languages, which stimulates intellectual/cognitive development.

(2) It is apparent that principals and program directors, as well as teachers, have benefited from a substantial investment in **staff development**, especially in matters pertaining to dual language program theory, the findings of research on dual language program implementation and effectiveness issues, and on improved teaching techniques. All of the staff of these dual language programs have received significant amounts of training in SIOP/TWIOP (Sheltered instruction/Two-way) strategies, and four schools have literacy coaches trained in dual language techniques who serve as ongoing in-house staff developers. NC DPI-sponsored training and conferences have been very important for all of these programs for professional development and networking. This is an important and necessary component for successful implementation of this powerful educational reform in school districts who wish to use it.

(3) **Administrative support** for this program is very strong for almost all of these NC schools, at the school level, the central administrative level, and the state level. This is crucial to success of the program. To monitor program implementation, most of the principals of these schools understand the model and know how to provide teachers with the kind of support that is needed, because it is a challenging teaching assignment. Several of these schools have dual language program coordinators who serve the role of administering the program and provide guidance in fidelity to implementation of program features that make a significant difference in success of the program.

(4) The greatest challenge for these dual language programs is **finding qualified bilingual teachers**, who are academically proficient in the language(s) of instruction, and who are certified to teach the mainstream curriculum for their Grade level. A second challenge is **finding or developing curricular materials in the non-English language** that match NC standards for each Grade level. Curricular materials in Spanish are more readily available, especially for the lower elementary Grades, but with each succeeding Grade level, finding appropriate materials is more challenging, and materials and resources in Mandarin Chinese have been a serious challenge for the Chinese teachers.

(5) **Parental support** is very strong for most of these schools. Every dual language program has a waiting list for the English speakers hoping to attend, and eight of the ten schools enroll English-speaking students by lottery because they have such a long waiting list. For most of these schools, Hispanic and Chinese parents have become strong advocates of this program model with each additional year of maturity of the program. While newly arrived immigrants often do not understand the purpose of bilingual schooling, several of the NC dual language schools have developed a growing parent advocacy
group to assist the newly arriving families. Three of the Spanish-English schools have had to work hard to keep up their Hispanic enrollment, because they are not located in a neighborhood with larger Hispanic demographics. Since English learners clearly benefit from these dual language classes, as demonstrated in our findings in Part III of this study, it is important to provide this information to Spanish-speaking families living in the attendance zones of these schools with dual language programs, and to expand these dual language services to more NC schools when possible.

Part III—Comparisons of Dual Language to Non-Dual Language Programs—Findings and Discussion

(1) Many dual language schools in Grades 3-5 have “strands” that define two groups of students—those students who receive dual language instruction in dual language classes, and students in the same Grades but in different classes not receiving dual language instruction. A third group of students is found in other schools that have no dual language classrooms and thus offer no dual language instruction. In Grades 6-8, all dual language middle schools in this study follow the “whole school model” (all students in the school receive dual language instruction), and other middle schools offer non-dual language instruction.

In Grades 3-5, students in dual language (DL) classrooms score significantly higher than both non-DL students in the same schools as the DL classes and higher than non-DL students in other schools with no DL classrooms. In Grades 6-8, the dual language students significantly outscore the non-DL students in other middle schools. These differences are large enough to be considered pragmatically and practically significant in terms of statistical effect size.

(2) Dual language average reading scores consistently exceed the statewide average reading score in Grades 3-8, while the average scores of non-DL students do not. The extra achievement increment afforded by participation in dual language instruction allows these students to significantly exceed on-grade-level instructional standards, as defined by the average state scores, while the non-DL students are at or slightly below these standards. In fact, the DL students in the upper grades (Grades 5-7) meet or exceed the average scores of students a year ahead of them in school.

While this phenomenon should be followed and verified longitudinally in later years, it is completely consistent with the theoretically-validated power of dual language programs to produce achievement gains that build cumulatively upon each other over 4-6 years. In other words, dual language programs typically exhibit smaller gains in the early Grades but produce larger gains as the cognitive demand of instruction increases after Grade 4, when non-dual language students (not receiving the increased cognitive development afforded by DL instruction) typically begin to fall behind. This finding is also completely consistent with the findings of other large research studies that have longitudinally followed dual language students over 5-6 years and have found relatively smaller gains in Grades K-3 and relatively larger gains in Grades 4-6. This consistent finding underscores the importance of following the results of dual language programs over 5-6 years at least, and not restricting policy making to short-term results from 2-3 year studies.

(3) The following groups of students were designated as defined student groups of interest:

(a) Current LEP students;
(b) Former LEP students;
(c) Language minority students not classified as LEP;
(d) Native-English speakers who are African-American;
(e) Native-English speakers who are Caucasian;
(f) Native-English speakers who are of other racial/ethnic background and are not language minority.
After analyzing the data separately for these defined student groups, and after comparing each group’s performance in dual language instruction to non-dual language instruction, we find that no group suffers from receiving dual language instruction and some groups benefit substantially. In addition, all dual language groups receive the substantial benefit of eventual deep proficiency in two languages rather than only one. In particular, we found that current LEPs benefit greatly from dual language instruction. Former LEPs (who have already “graduated” from a previous English learner program such as English-as-a-Second-Language) also benefit, but to a lesser degree than current LEPs. Students who are language minority but not LEP benefit similarly to former LEPs, receiving small-to-moderate achievement increases.

We found that African-American native-English speakers benefit greatly from participating in dual language classes. In fact, the scores of African-American students in dual language instruction are at or near the statewide average scores for each Grade. African-American students overall were identified in Part I of this study with an even-larger-than-expected (compared to other research) achievement gap, when compared to Caucasian native-English speakers. These substantial advantages accruing to African-American students receiving dual language instruction were not completely unexpected, given the authors’ prior research with Houston, Texas Independent School District’s 200,000+ students, one-third of whom are African-American. However, we were pleasantly surprised that NC African-American students in dual language, even with substantial numbers of low-socio-economic status (based on statewide statistics), responded with significantly higher Reading achievement scores than either of the other two groups of African-Americans in non-dual language instruction.

(4) The group of native-English speakers of other (non-Caucasian, non-African-American) racial/ethnic background was mostly too small for reliable sub-group analyses, and so no findings for decision-makers are presented here for this group. It is worth noting that the Other students in dual language classes numerically scored at or above the Other students in non-dual language classes in every Grade, despite the small group sizes.

(5) Caucasian native-English speakers, while demonstrating the highest achievement scores of any of the groups, also receive small-to-moderate increases in Reading achievement scores when they participate in dual language classes. In fact, they score numerically higher than Caucasians in non-DL instruction in every Grade, and do so significantly in Grades 3, 5, and 6. There is no evidence that their academic development (as measured by the EOG Reading test in each Grade) is hurt in any way, and they are, as are all DL students, experiencing the advantages of acquiring full proficiency in English plus a second language (e.g., Spanish, Chinese). For Caucasian native-English speakers, a dual language program functions like a gifted program, providing everything available in the regular instructional program, plus more.
Overall Summary

The above findings for African-American students imply strongly that addressing the existing African-American achievement gap in North Carolina might offer a strong justification for further developing and expanding dual language instruction by itself. In addition, as previously noted, there are very large potential benefits of expansion and further development of dual language programs for language minority students, especially LEPs. In combination, the prospect of substantially raising the test scores of these two most-at-risk groups offers a powerful incentive to consider dual language instruction as the “schooling of choice” for many more North Carolina students. When this prospect is combined with effectively offering a “gifted” program for the Caucasian native-English speakers who can benefit from it, and also potentially addressing the instructional needs of Title I students of all racial/ethnic backgrounds (especially large groups of Caucasians and African-Americans), the reform called dual language education becomes attractive as a “win-win-win” scenario for North Carolina students whose school districts are interested in this program model.

Since the data sets analyzed in this study are cross-sectional (2008 testing only with other databases from 2006-2009) and since the analyses are statistically descriptive and do not yet control (statistically or experimentally) for extraneous variables that could theoretically affect the results to some degree, these findings must be considered as indicative and suggestive rather than confirmed at this point. However, the findings of this study are especially suggestive for the following reasons:

(a) They consistently favor dual language instruction across all groups and situations analyzed, rather than only some.

(b) The effect sizes observed with dual language instruction are consistent with effects found in other large-scale research studies.

(c) Overall, the study’s sample sizes are large enough for meaningful analysis of both groups and subgroups of interest. Such analysis is avoided for groups that are too small to support it.

(d) The analyses of significant differences use statistically conservative criteria of practical and actionable significance and avoid situations where overly large or overly small group sample sizes might produce spurious statistical significance (Type I error) or spurious lack of statistical significance (Type II error). In other words, the real effects of dual language instruction may well be larger than indicated in this study.

(e) The groups in this study most positively affected by DL classes are those that educational and psychological theory says should be most affected by a cognitively stimulating educational program. These are the English learners, those who are known to be educationally at-risk and who need additional cognitive development to achieve gap closure, and those in Grades 4-6. These are the grades in which extra cognitive development is most needed to sustain prior achievement levels as the cognitive demand of the instructional curriculum begins to increase after Grade 3, and continues to do so into the middle and high school years.

Considering the findings from Part I, Part II, and Part III, some very good things appear to be happening in North Carolina schools with dual language programs. Dual language teachers and administrators are intensely focused on diagnosing and meeting student needs and providing extra cognitive stimulation in doing so. Student achievement is increasing in dual language classrooms, and all participating students are acquiring two languages rather than just one. Based on findings in Parts I, II, and III, one can easily
make the case that this innovation should be encouraged, developed further, studied longitudinally, and supported in North Carolina school districts that are interested in using this innovation to reform their instruction for all students.

References


