

# Assessing the Effectiveness of Tennessee's Pre-Kindergarten Program: First Interim Report

December 12, 2007

---



995 Goodale Blvd • Columbus, Ohio 43212

Phone: (614) 220-8860 • Fax: (614) 220-8845  
[www.strategicresearchgroup.com](http://www.strategicresearchgroup.com)



The **Strategic Research Group (SRG)** is a full-service research firm that provides data collection, consultative, and research services. SRG specializes in conducting public opinion surveys, program evaluations, policy assessments, customer satisfaction studies, and community needs assessments on national, state, and local levels.

Comptroller of the Treasury, Office of Research. Authorization Number 307356, 50 copies, February 2008. This public document was promulgated at a cost of \$1.89 per copy.

## Foreword and Acknowledgements

This **First Interim Report**, produced under contract with the Tennessee State Comptroller's Office, provides the initial results of an analysis of student outcomes for the first five cohorts of students who participated in Tennessee's Pre-K program between 1998-1999 and 2002-2003. Statistical analyses explored the short- and long-term impact of Pre-K participation on student assessments in Kindergarten, First, Second, Third, and Fourth Grades. This report also provides a descriptive overview of Pre-Kindergarten students and programs in Tennessee (see Appendix A) and offers a detailed overview of the research methodology used to assess the effectiveness of Pre-Kindergarten programs in Tennessee.

This report has been made possible due to the support from the Tennessee Department of Education (TDOE) and the Tennessee State Comptroller's Office of Education Accountability. In particular, we wish to thank the following individuals for their expertise and assistance in obtaining the data necessary for analysis, providing important background information, and for their ongoing support of this research:

**Connie Casha**, Director of Early Childhood Programs for the Office of Early Learning

**Corey Chatis**, Director of Data Quality

**Dr. Phillip Doss**, Assistant Director for the Tennessee State Comptroller's Office of Education Accountability

**Dan Long**, Senior Executive Director for the TDOE Office of Assessment, Evaluation and Research

**Bobbi Lussier**, Executive Director for the Office of Early Learning

# TABLE OF CONTENTS

---

**Executive Summary** ..... 5

**Evaluating the Effectiveness of Tennessee’s Pre-Kindergarten Program** ..... 8

**Background and Objectives of the Present Study** ..... 9

*Organization of This Report* ..... 10

**Methodology** ..... 11

*Data Sources* ..... 11

**1. Pre-Kindergarten Demographic File** ..... 11

**2. Education Information System Data** ..... 12

**3. K-12 Student Assessment Data** ..... 12

*Data Management* ..... 15

*Sampling Strategy* ..... 20

*Analytic Approach* ..... 22

**Results** ..... 22

**Discussion and Next Steps** ..... 38

    Appendix A. Background and Implementation of Tennessee’s Pre-Kindergarten Program ... 42

**Pre-K Pilot Sites** ..... 43

**Student Eligibility** ..... 43

**Classroom Requirements** ..... 44

**Program Effectiveness** ..... 45

    Appendix B. Identifying Discrepancies in Gender and Race in Student Assessment Records ..... 46

    Appendix C. Imputing Missing Demographic Information ..... 47

    Appendix D. Pre-K Group Demographic Profile for Each Grade/Year ..... 48

## Executive Summary

Since 1999, Tennessee’s Pre-K program has experienced significant growth, with an increasing number of “at-risk” students participating every year. In 1998, the Pilot program began with 600 three- and four-year olds served. Projected enrollment for the 2007 academic year is 17,000.

With ten cohorts of children having participated in the state Pre-K program as of the 2007-2008 academic year, a critical question concerns the impact of the program on assessment scores as students enter and progress through the public school system. Preliminary analyses conducted by the State of Tennessee suggest that Pre-K participants fare better on standardized assessments relative to other at-risk students who did not participate in the Pre-K program. However, the present research uses rigorous sampling techniques and statistical analysis to empirically test the hypothesis that Pre-K participants perform better on standardized assessments in Kindergarten through Fifth Grade than a comparable group of at-risk students who did not participate in the Pre-K program.

Tests of statistical significance were conducted for five cohorts of students within Kindergarten, First, Second, Third, and Fourth Grade comparing the results of standardized assessments for students who participated in Pre-K to the results for a similar group of at-risk students who did not participate in Tennessee’s Pre-K program. “At-risk” was defined for purposes of this study as receiving free/reduced price lunch (FRPL) during at least one school year in the timeframe under study (1999-2004). The table below shows the grade levels analyzed for the First Interim Report.

### Cohort Trajectories and the School Years Covered in the First Interim Report

	1999-2000	2000-2001	2001-2002	2002-2003	2003-2004
Cohort 1	K	1st	2nd	3rd	4th
Cohort 2		K	1st	2nd	3rd
Cohort 3			K	1st	2nd
Cohort 4				K	1st
Cohort 5					K

Many Pre-K students attended schools that did not administer assessments in Grades K-2. Further, some inaccuracies and inconsistencies in the data resulted in excluding some cases from the analysis. Thus, due to the resulting small number of Pre-K participants for whom data were available across multiple years, longitudinal tests (i.e., analyses using data from the same students measured at multiple points in time) could not be conducted for these cohorts for the time period studied in this report. However, rigorous sampling techniques were used to construct valid comparison groups of at-risk students who did not participate in Tennessee’s Pre-K program. Specifically, the non-Pre-K samples were selected such that they mirror the Pre-K groups with regard to school district, gender, race, and FRPL status. Analysis of Variance (ANOVA) was used to examine the relative impact of Pre-K participation including differences based on race and gender on assessment scores between 1999-2000 and 2003-2004.

Across analyses of these five cohorts of students over five years, there were numerous positive effects associated with Pre-K participation such that **there do seem to be some short-term and long-term benefits of Pre-K participation on assessments in Grades K-4**. However, the pattern of results generally seems to change between cohorts. For some cohorts in some years, Pre-K students scored better than the comparison group of at-risk students who did not participate in Pre-K. The effects are seen in many assessments, including reading, language, word analysis, spelling, social studies, and science—although the benefits do not appear to be consistent for most of the assessments. One notable exception, though, is that beneficial effects associated with the Pre-K program are consistently evident in reading and language arts.

For assessments of reading and language arts (both NRT and CRT assessments) there is a recurrent pattern of results. Specifically, **non-white females who attended Pre-K scored higher than non-white females who did not attend Pre-K**. Further, among students who participated in Pre-K, non-white female students scored comparably to white female students. This pattern was not observed for non-white males, for white males, or for white females. This is an intriguing finding because it suggests that the non-white female students, in the aggregate, reap the greatest benefits of Pre-K participation, particularly in the area of reading and language arts. Further, this pattern of results emerged over multiple cohorts (Cohorts 1, 3, 4, and 5), over multiple assessments (reading, language, vocabulary, and word analysis), and in multiple grades (K, 1, 3, and 4). This suggests that Pre-K participation may help to close the “achievement gap” that is often observed between white and non-white students—but the impact is localized such that non-white females, on average, tend to show the greatest benefits of Pre-K participation, at least with regard to certain assessments. Further exploration of this finding will be necessary as the evaluation progresses.

For assessments involving mathematics, the results are less consistent but do reveal a positive impact of Pre-K participation, depending on grade level. For example, among Cohort 1 students in the Third Grade, Pre-K participation was associated with improved mathematics performance among non-white students relative to a comparison group of at-risk non-white students who did not participate in Pre-K—but this effect was not observed at the next grade level. Among Cohort 2 students, there was a significant difference in mathematics scores in First Grade such that Pre-K students scored higher on this assessment compared with non-Pre-K students, but again, this pattern was not evident in other grade levels for this cohort. A more complex pattern was observed for Cohort 3: although Pre-K participation appears to be associated with higher mathematics performance in Kindergarten and First Grade for non-white students, it is also associated with lower performance for male students. Considering these results together, **the impact of Pre-K participation on mathematics performance is variable and suggests a limited short-term impact** based on the analyses conducted to date.

Likewise, the benefits of Pre-K participation are limited in terms of science and social studies—these effects vary by cohort and grade level. There is some evidence to suggest that Pre-K participation is associated with higher performance on the science and social studies assessments, but again the results are inconsistent.

Further analyses are planned and this evaluation is ongoing. Many of the preliminary findings in this report must be explored further. The Second Interim Report will compare student achievement between Pre-K and non-Pre-K students between the years 2004-2005 through 2006-2007. Over time, we anticipate the challenges experienced for this report in terms of small sample sizes will be attenuated. This is because at higher grade levels (Third, Fourth, and Fifth Grades), all students are required to complete standardized tests. In the data to be analyzed for the Second Interim Report, five of the eight cohorts to be studied will be in the Third, Fourth, or

Fifth Grades. We anticipate that for these grades, longitudinal analysis may be possible given that assessments in Grades Three through Five are mandatory and thus there will be larger numbers of students who can be followed over time. Also, the Pre-K program has expanded in recent years, resulting in more Pre-K student outcome data to analyze. Additionally, because Education Information System (EIS) data are available for two of the three years covered in the Second Interim Report, we expect to compare Pre-K and non-Pre-K students with regard to other indicators of academic achievement, such as attendance and disciplinary issues. The inclusion of multiple outcomes will offer a more thorough examination of the effectiveness of Tennessee's Pre-K program that was not possible for this First Interim Report.

# Evaluating the Effectiveness of Tennessee’s Pre-Kindergarten Program

The State of Tennessee has commissioned an evaluation of the effectiveness of its Pre-Kindergarten program. The primary objective of the project is to assess whether children who attended a state-funded Pre-Kindergarten program perform better academically in the short and long term than a comparable group of peers who did not attend Tennessee’s Pre-Kindergarten program. The evaluation will also investigate whether various characteristics of Tennessee’s Pre-Kindergarten programs impact short- and long-term achievement among students who attended these programs.

Over the three years of this evaluation, eleven cohorts of students will be tracked through the Fifth Grade in order to capture both short- and long-term academic achievement; this process will begin with the cohort of students who attended Pre-Kindergarten in the 1998-1999 academic year. There will be a total of five cohorts that have assessment data spanning from Pre-Kindergarten to Fifth Grade, and data for at least one grade for the other six cohorts.<sup>1</sup> The results from analyses of these cohorts will be covered over the course of six project reports (two annual, three interim, and one final report).

The objective of this First Interim Report is to provide the Office of Education Accountability (OEA) and the General Assembly a report on the initial results of an analysis of student outcomes for the first five cohorts of students who participated in Tennessee’s Pre-K program between 1998-1999 and 2002-2003. Table 1 summarizes the time period under study for this report.

**Table 1. Cohort Trajectories and the School Years Covered in this Report**

	1999-2000	2000-2001	2001-2002	2002-2003	2003-2004
Cohort 1	K	1st	2nd	3rd	4th
Cohort 2		K	1st	2nd	3rd
Cohort 3			K	1st	2nd
Cohort 4				K	1st
Cohort 5					K

Statistical analyses explore the short- and long-term impact of Pre-K participation on student assessments in Kindergarten, First, Second, Third, and Fourth Grades. This report offers a detailed overview of the research methodology used to assess the effectiveness of Pre-Kindergarten programs in Tennessee and also provides a descriptive overview of Pre-Kindergarten students and programs in Tennessee (see Appendix A).

---

<sup>1</sup> This does not imply that the analysis will include data for all students for all years, as many students are not assessed in Kindergarten, First, and Second grade.

## Background and Objectives of the Present Study

### The TDOE Study

In Spring 2006, the Tennessee Department of Education (TDOE) examined student outcomes from the first four cohorts of students who participated in the Pilot Pre-K program. Preliminary data analysis conducted by TDOE on assessment results for students who participated in the Pilot Pre-K program suggests that at-risk students who participate in Tennessee Pre-K perform better on standardized achievement tests than the at-risk students with no Pre-K experience. The cohorts, or groups of students studied, are described in Table 2.

**Table 2. Cohorts studied in Preliminary Analysis by Tennessee Department of Education**

	1998-1999	1999-2000	2000-2001	2001-2002	2002-2003	2003-2004	2004-2005	2005-2006
Cohort 1	Pre-K	K	1st	<b>2nd</b>	<b>3rd</b>	<b>4th</b>	<b>5th</b>	
Cohort 2		Pre-K	K	<b>1st</b>	<b>2nd</b>	<b>3rd</b>	<b>4th</b>	5th
Cohort 3			Pre-K	K	<b>1st</b>	<b>2nd</b>	<b>3rd</b>	4th
Cohort 4				Pre-K	K	<b>1st</b>	2nd	3rd

Note: Grade levels in bold type were studied by TDOE.

Assessment data for students who participated in the Pilot Pre-K program in 1998-1999, 1999-2000, 2000-2001, and 2001-2002 were analyzed by TDOE to determine whether there were any differences between Pre-K participants and their at-risk peers (identified by free/reduced price lunch status) in assessments completed in First through Fifth Grade.

The results of the TDOE analysis on students' reading and mathematics scores suggest that overall Pre-K participants scored slightly higher on norm-referenced assessments (NRTs) in these subjects in First, Second, Third and Fourth Grades. Also, the trend was for Pre-K participants to score slightly higher on criterion-referenced assessments (CRTs) in these subjects in the Third, Fourth, and Fifth Grades. Generally speaking, Pilot Pre-K participants' scores were slightly higher than other at-risk students but not quite equivalent to students not identified as "at-risk." This suggests that participation in Pre-K may help to close the "achievement gap" that tends to persist nationally between at-risk and non-at-risk students.<sup>2</sup>

A summary of the results is available from the TDOE website (see Footnote 2). Although these results are quite encouraging, several important questions are left unanswered. First and foremost, although students who participated in Pre-K appear to score higher (on average) on assessments compared to other students identified as being "at risk," it is unclear from these results whether the differences are statistically significant—that is, it is difficult to say whether the differences observed between these groups are due to naturally occurring variation between students or due to some systematic difference in student performance.

<sup>2</sup> An overview of the TDOE analysis is available for download from the Internet at: [http://www.tennessee.gov/education/prek/doc/prek\\_in\\_tn.pdf](http://www.tennessee.gov/education/prek/doc/prek_in_tn.pdf).

Further, it is unclear whether the sample of Pre-K students (those for whom assessment results are available in Kindergarten, First, Second, Third, and/or Fourth Grades) has been compared to a **comparable** sample of at-risk students who did not participate in Pre-K. That is, it is difficult to say to what extent the comparison groups resemble the Pre-K groups in terms of gender, race, or school districts attended. In addition, it is not clear how a student's status for being "at risk" is defined in the TDOE analysis. Traditionally FRPL status is used as a proxy for socioeconomic status, and thus FRPL status is used to determine whether a student is "at risk;" however, because a student's FRPL status may change once or multiple times in any given timeframe, exactly how this indicator is used to define risk may be open to some debate.

### **The SRG Study**

The present evaluation conducted by Strategic Research Group seeks to clarify and build on preliminary TDOE results in several ways. First, additional data is now available to validate these preliminary findings, including more students at more grade levels. Further, although TDOE has presented data from reading and mathematics assessments, the results of other assessments may also be examined to see whether Pre-K participation has an impact. Statistical analysis will determine whether group differences (comparing Pre-K participants and non-Pre-K participants) fall within a range to be expected due to chance, or whether they suggest a systematic difference between the groups. The goal of these analyses is to assess whether participation in Pre-K is associated with higher standardized test scores in the short term (Kindergarten through Second Grade) and in the long term (Grades Three through Five).

To address potential concerns about the preliminary TDOE analyses of Pre-K student outcomes—in particular making sure that the Pre-K program participants are compared to a similar sample of non-Pre-K attendees—SRG analysts have worked extensively with the data. Student records were checked thoroughly so that as many Pre-K participants as possible would be included in the analysis. In addition, through rigorous sampling techniques, we have constructed proportionally similar groups of at-risk students who did not attend Tennessee's Voluntary Pre-K program. Comparative analyses will use a reference group that is as similar as possible to the Pre-K participants in terms of gender, race, school system attended, and FRLP status.

This research compares both the short-term (Grades K-2) and long-term (Grades 3-5) effects of Pre-K participation, affording a more comprehensive look at student outcomes. Further, in these analyses SRG explores the relative impact of Pre-K participation on student subgroups—in particular, whether differential effects of Pre-K participation are found as a function of gender or race.

Finally, as part of this evaluation SRG will explore the relative impact of various program characteristics to see whether there are differences in student outcomes as a result of program structure or implementation. Such analysis is not possible for the First Interim Report because for the time period under study all students participated in the same fundamental program (the Pilot Pre-K program). Subsequent reports will explore this question as more data becomes available.

### ***Organization of This Report***

This report provides a detailed review of the methodology and results of an analysis of student assessment results from 1999-2000 to 2003-2004, with the following sections:

- The **Methodology** section provides an overview of the data sources used in this investigation, SRG's management of the data, and how data sources were combined and prepared for analysis. Further, this section provides the details of our sampling strategy, explains how the comparison groups of non-Pre-K students were constructed, and offers an overview of our analytic approach.
- The **Results** section provides a detailed summary of statistically significant findings associated with participation in the Pre-K program arranged by Cohort and Grade Level.
- The **Conclusions** section provides a general summary of the results along with next steps for the analysis and important questions to be addressed in the Second Interim Report, due in March 2008.
- In the **Appendices**, we provide an overview of Tennessee's Pre-K program, examples to illustrate how we resolved discrepancies or irregularities between data sources, and a descriptive profile of the Pre-K groups for each grade and year examined in this report.

## Methodology

For the present study, TDOE provided the following datasets: student assessment data from 1999-2000 through 2004-2005, a list of Pre-K attendees each year since 1998, and student demographic information from TDOE's Education Information System (EIS). Great care was taken by TDOE and SRG to ensure student anonymity. No identifying information was provided along with student outcome data. To conduct the present study, these data sources were merged, and any irregularities or inconsistencies between the sources had to be addressed and reconciled. The Methodology section of this report details the nature of each data source, how SRG combined them to construct samples of Pre-K and non-Pre-K participants, and how this process impacted our analytic approach.

## Data Sources

For the First Interim Report, SRG primarily drew from two data sources: 1) Pre-Kindergarten demographic data, and 2) K-12 student assessment data. SRG was also provided Education Information System (EIS) student demographic data from the 2005-2006 school year. Because the current phase of the evaluation spans from 1999-2000 through 2003-2004, EIS data were used primarily to verify discrepant information in other data sources.

### 1. Pre-Kindergarten Demographic File

The Pre-Kindergarten (Pre-K) demographic file is a database maintained by the TDOE's Office of Early Learning. It was provided to SRG via the Director of Data Quality for the TDOE. The database spans eight academic years from 1998-1999 to 2005-2006. Starting with the 2006-2007 school year, information about Pre-K students is included in the Education Information System (for more information about the EIS, see the following section).

The Pre-K database contains information on the school (including county, system/LEA, and school/provider name), program information (e.g., Pre-K funding source), and student demographic information (date of birth, gender, race, free/reduced price lunch (FRPL) status, special education status, whether English is the student's native language, and whether the school provided transportation). Although information is not available for all

variables for all years in the Pre-K demographic file, the most important function of this data source is to identify students who participated in Tennessee's Voluntary Pre-K Program beginning in 1998-1999 through 2005-2006.

To protect student confidentiality and to comply with federal regulations regarding student FRPL status, SRG did not obtain student names or Social Security Numbers. Social security numbers, however, were encrypted by TDOE so that the various data sources could be combined for the data analysis. This permitted SRG to link student assessment results with student demographic information and Pre-K participation data, but in a way that maintained student confidentiality.

## **2. Education Information System Data**

The Education Information System (EIS) is a web-based data repository containing detailed student, teacher, school, and district level information. All schools input student information in a standardized format, and the EIS system is designed to catch data entry errors. EIS data are available beginning with the 2005-2006 school year. Although EIS includes data for prior school years, SRG was informed that these data are not complete and the state-assigned student ID number was only implemented in 2005-2006.

EIS files were provided to SRG by the TDOE Director of Data Quality. The data are in the form of spreadsheets that include demographic information, attendance records, disciplinary records, and special education records. EIS contains data for students in Kindergarten through Twelfth Grade, and for Pre-K students beginning in 2006-2007.<sup>3</sup>

Student demographic information contained in the EIS is more complete than the Pre-K demographic file because not all demographic variables were collected in the Pre-K records for all years. This is also the case with the assessment data. However, because the time period under investigation pre-dates the information in the EIS, and because student demographic information is available in the student assessment data (see discussion of this source below), the primary purpose of the EIS at this phase in the evaluation is to verify discrepant information in other data sources.

## **3. K-12 Student Assessment Data**

The third data source available for this evaluation is standardized assessment scores for students. These files were provided to us by the TDOE Assessment, Evaluation, and Research Division via the Department's Director of Data Quality. SRG requested and received scores for the 1999-2000 through 2005-2006 school years. SRG will obtain scores for future years as they become available and necessary for subsequent stages in the analysis. The files contain: 1) demographic characteristics of students and 2) test scores in the following subject areas: reading/language arts, mathematics, science, and social studies, along with composite scores by academic year.

The TCAP (Tennessee Comprehensive Assessment Program) is the principal tool for assessing the performance of public school students in the State of Tennessee. The TCAP includes Tennessee-specific assessments which allow students, parents, and educators to interpret test scores as they relate to Tennessee's state curriculum standards.

---

<sup>3</sup> SRG did not obtain data for students in Grades 7-12 as they are not needed for the present evaluation.

For students in Grades K-2, TCAP currently consists of Norm-Referenced Tests (NRT). Students in Grades 3-8 currently take Criterion Referenced Tests (CRT). NRTs measure student performance relative to other test takers. Comparatively, CRTs measure performance according to specific standards, and test items are directly linked to specific performance indicators in the state curriculum. The TCAP Achievement test is mandated for all students in Grades 3-8. The test is not mandated for Grades K-2; however, school systems may elect to test students in Grades K, 1 and/or 2, and their choice to test may vary from year to year.

The test for Kindergarteners includes reading, language arts, and mathematics. At Grade 1, the test includes reading, language arts, mathematics, science, social studies, word analysis, vocabulary, and math computation. The Grade 2 test includes all these subjects and also incorporates spelling. The CRT assessments for Grades 3-8 include four subject areas: reading/language arts, mathematics, science, and social studies. Tennessee students are assessed each spring.<sup>4</sup>

### **Comparability of NRTs and CRTs**

Although both NRTs and CRTs are important and valuable in their use and application, there are some issues in terms of their comparability. For example, when CRTs are employed, each individual student's results are compared with a predetermined standard. The performance of other students who also took the test at the same time is not taken into consideration in evaluating the results. Student scores are typically reported in terms of the number of items correct, or the percentage correct. In contrast, for NRTs, each individual student is compared with other students who took the test, and the score reflects that student's performance relative to other students (not a predetermined criterion). Scores are typically reported in terms of a percentile or stanine, which indicates the student's position relative to a national sample of other test-takers in the same cohort.

### **Changes in Assessments and Implications for Analysis**

The present report includes an analysis of student outcomes between 1999–2004, five academic years. During this timeframe, a number of changes took place in Tennessee with respect to student assessment, and these changes have important implications for the types of analyses that can be conducted to assess student outcomes. The first change of significance is the adoption of CRTs in Grades Three through Eight.

From 1999-2000 through 2002-03, student assessments were NRT tests created by CTB Terra Nova. The data for these years are all norm-referenced and have used the same scale (i.e., same minimum and maximum scores) since 1998. There were no cut points for NRT data to indicate proficiency levels during this time; proficiency levels (advanced, proficient, below proficient) were introduced with criterion-referenced tests.

Of particular importance to the present study is that prior to 2003-2004, Tennessee used a norm-referenced test for students in Grades Three through Eight for reading, language arts, mathematics, science and social studies. The implication of the change to a CRT is that the ability to compare student outcomes across grade levels is problematic because the analysis depends upon the comparability of the measures. Because there are significant

---

<sup>4</sup> Note: The scope of the present analysis is focused on student performance in grades K-5.

conceptual and practical differences in the nature of the CRT and NRT assessments, longitudinal analyses across these measures are not feasible.

Comparatively, there has been no change in the K-2 assessments between 1999 and the present. Students in Grades K-2 take norm-referenced tests, and Terra Nova NRT tests have been used for the entire period of study in the present analysis. However, individual school systems may choose whether or not to administer these tests, and according to data provided by TDOE, a variable proportion of districts elect to administer these examinations each year. Thus our ability to conduct longitudinal analyses across years depends on the availability of the data—that is, whether enough Pre-K students attended schools in which they took an assessment in Kindergarten, First, and Second Grade, and whether the tests were given in consecutive years to permit analysis over time.

TDOE provided SRG a spreadsheet summarizing the number of assessments administered in Grades K-2 by each Local Education Agency (LEA) each year between 1998 and 2006. Only 5 of 137 LEAs continuously participated in assessments for Grades K-2 during this time period. LEAs may administer tests for one, two, or three of these grade levels in a given year, and they may change their decision to administer assessments each year. Thus, there is a great deal of variability in the number of schools administering assessments for students in Grades K-2 across this time period.

Using the information provided by TDOE, SRG estimated the percentage of school systems administering tests each year, and these estimates are summarized in Table 3. These are considered estimates because the information provided does not conclusively state whether or not tests were administered to all students (some systems may test only some of their students or administer a small number of out-of-level tests). Still, it is clear that less than 25% of systems typically administer assessments in Kindergarten, about half administer tests in First Grade, and most school systems administer tests in Second Grade.

**Table 3. Estimated Percentage of Tennessee School Systems Administering Assessment Tests in Grades K-2 between 1999-2000 and 2003-2004**

	Grade K	Grade 1	Grade 2
1999-2000	20%	48%	83%
2000-2001	23%	56%	90%
2001-2002	19%	55%	90%
2002-2003	23%	63%	91%
2003-2004	24%	61%	87%

The variability in the number of students being tested each year has important implications for the present study. Because overall very few students are tested in Kindergarten, only about 20% of the Pre-K students should have Kindergarten assessment scores (assuming that there is no bias in the way Pre-K students are distributed across school systems, which is likely not the case). True longitudinal analysis requires that the same students are assessed at multiple points in time, and each student must be measured at each point in time. Thus, to be able to conduct longitudinal analysis over Grades K-2, it is necessary to have measurements for each student at all three time periods. Because so few students have assessment results in Kindergarten, this is a serious constraint that severely limits the

number of students who would be included in such an analysis (more information is provided on the number of Pre-K students for whom data are available later in this report; see Table 4 on page 19). Therefore the present analyses compare Pre-K participants to a similar group of non-Pre-K participants in the same cohort and grade level.

With this in mind, SRG began the process of identifying Pre-K students, locating their assessment results, resolving any data discrepancies or inconsistencies, and drawing a comparable sample of at-risk non-Pre-K attendees.

## ***Data Management***

As was mentioned previously, SRG requested and received assessment data for the 1999-2000 through 2004-2005 school years. The data were provided in two files: one containing the scores for the norm-referenced assessments (administered to students in all grades until 2001-2002), and the other containing the scores for the criterion-referenced assessments (administered to students in Grades Three through Five starting in 2002-2003). In the original datasets that were provided by the TDOE, there were 507,501 cases in the NRT dataset and 346,040 in the CRT dataset.

The datasets were reorganized so that there was one dataset for each cohort to be analyzed for this report. For example, the Cohort 1 dataset is comprised of students in Kindergarten in 1999-2000 through Fourth Grade in 2003-2004. The Cohort 2 dataset is comprised of students in Kindergarten in 2000-2001 through Third Grade in 2003-2004, and so on. Although this report does not include the 2004-2005 school year, data from this school year were used in this data management phase to help identify potentially erroneous records and in the imputation of demographic information for students missing information in other grade/years (discussed below).

Several steps were required in order to prepare the assessment data for analysis:

### **1. Identify and exclude records with duplicate ESSNs.**

The first step in preparing the data for analysis was to identify and exclude records with duplicate ESSNs. Each grade/year dataset contained a small number of cases with duplicate ESSNs, meaning that there were two (and in a very small number of instances, three) sets of scores for the same grade level and school year linked to the same ESSN. An examination of duplicate records found that in most cases, although the ESSN was the same, the student's demographic information was not, indicating that the assessment scores were for different students. When possible, an effort was made to cross-check the demographic information linked to the ESSN with the EIS data to determine which record was incorrect. In particular, this process was done for the Kindergarteners for most years, as these were the smallest datasets and contained the smallest number of students who had attended Pre-K, thus it was most important to attempt to retain as many valid records as possible. For other grade/years, because of the relatively large number of cases, it was not feasible to cross-check all duplicate records individually against the EIS. In these cases both records were excluded from analysis. It should be noted however, that for most grade/years, cases with duplicate ESSNs represented a very small proportion of all cases for that grade/year.

### **2. Identify and flag records for students with assessments scores for more than one grade level in the same school year.**

The second step was to identify and flag records for students that had assessment scores for more than one grade level in the same school year. For example, if a student had one set of scores as a first grader in 2001-2002 and another set of scores in that same school year as a second grader, that student's record was flagged and excluded from all subsequent analyses. Although it is reasonable for a student to have scores at the same grade level for consecutive years (e.g., scores as a first grader in both 1999-2000 and 2000-2001) as a result of grade retention, multiple sets of scores in the same school year at different grade levels is indicative of an error.<sup>5</sup> An examination of a number of these instances found that in each instance, the two sets of scores, while being linked to the same ESSN, demographic information (such as date of birth, gender, and/or race) differed. It should be noted that these instances occurred very rarely, ranging from 15 cases in 2000-2001 to 391 cases in 2003-2004.

### **3. Examine the consistency of gender and race across students' available set of scores in each grade/year.**

Next, SRG researchers checked the consistency of gender and race across students' available set of scores in each grade/year. For example, for a student who had scores as a Kindergartener in 1999-2000 through Fifth Grade in 2004-2005, we counted the number of sets of scores for which the student was identified as being of each category of gender and race. The ideal circumstance, of course, is one in which the student's reported gender and race were the same for each grade/year. This was the case for the great majority of cases. Instances in which gender and/or race values were different, however, needed to be investigated further.

There are two likely causes for discrepancies in gender and race across grade/years. First, values for these characteristics may have been reported or entered incorrectly at one or more points in time. Second, the Social Security Number may have been reported or entered incorrectly for a student in one or more grade/years. For these discrepant records, two approaches could be taken: first, exclude from analysis *all* sets of scores for that ESSN, or two, exclude those most likely to contain incorrect information. SRG believes that in general a conservative approach is the best approach; that is, we feel it is better to exclude a case with valid information from the analysis than include one with invalid information. That said, however, we also think it is important to make an effort to retain as many accurate records as possible. With this in mind, we devised a strategy for determining which records would be retained and which would be excluded from analysis.

For any students with discrepant values for gender and/or race across sets of scores, if one value did not appear at least three times more often than any other value(s) for that characteristic, all sets of scores for that student were flagged and excluded from analysis. If one value did appear at least three times more often than any other value(s), the scores for that value were retained as it was considered to be valid and only the set(s) of scores for the inconsistent value(s) was/were excluded. For examples of how this approach was implemented, please see Appendix B.

### **4. Identify students who were missing information on the key demographic variables of interest and impute information from other records when possible.**

---

<sup>5</sup> This was verified by the Senior Executive Director for the TDOE Office of Assessment, Evaluation and Research.

The next step was to identify students who were missing information on gender and race for one or more grade/years and to impute values from other grade/years when appropriate. It was important to conserve as much of this demographic information as possible so that fewer cases would be lost during the processes of sampling and analysis. Sampling and data analysis will be discussed in greater detail below; however, it is necessary to state here that these demographics were key factors in sample selection and were also important components of the data analysis. Because gender and race should remain constant over time (and any discrepancies indicate error, an issue that was resolved in the preceding step), if a student's gender or race was not reported in a grade/year but was reported in another, the student's value for the missing grade/year(s) can be imputed from other records containing the missing information for the student.

As discussed previously, however, it is important to ensure that the imputed value is a correct one. Thus, we used a strategy whereby a student's gender or race was imputed only if they had at least two other sets of assessment scores containing a value for gender or race—and, of course, it had to be the same value for all available sets of scores).<sup>6</sup> For examples of how this approach was implemented, please see Appendix C.

Among all of the grade/years included for this report, the percentage of students whose gender had not been reported did not exceed 2% (and was usually less than 1%). The one exception was the Kindergarten class 2002-2003, of which approximately 11% did not have a reported value for gender. After the imputation process, the occurrence of missing values was reduced to less than 0.5% for all grade/years except the aforementioned Kindergarten class of 2002-2003, in which only 0.6% of students were still missing gender after the imputation process.

Missing information was more common in the case of race. The percentage of students whose race had not been reported was less than 12% for 12 of the 15 grade/years covered in this report, and after the imputation process less than 5% of students were still missing on race for 11 of 15 of the grade/years. Of particular relevance, only 51% of Kindergarteners in 2000-2001 had a value for race and race was not reported at all for Kindergarteners in 1999-2000. The imputation process increased the percentage of students with a value for race from 51% to 88% for Kindergarteners in 2000-2001 and from 0% to 78% for Kindergarteners in 1999-2000.<sup>7</sup>

Next, we needed to determine how to deal with missing data for students' FRPL status. There were two issues in this regard. First, for some students in some years, FRPL status information was missing. Second, upon inspection of the data, it became clear that there was some variability in FRPL status. That is, a given student's FRPL status might change from one year to the next. Thus, for purposes of the data analysis, we had to make some decisions about how to determine whether a student should be considered "at risk" based on FRPL status over the time period under investigation.

After some consideration, we decided to consider a student "at risk" based on FRPL status if the student had been identified as receiving FRPL at least once in the years in the time

---

<sup>6</sup> Because Cohort 5 is comprised only of two grade/years (the Kindergarten class of 2003-2004 and the First Grade class of 2004-2005), no imputations could be made for demographic characteristics.

<sup>7</sup> Despite the fact that demographic variables could be imputed from other sources, at the conclusion of the data management phase, there were too few Pre-K students who were assessed in Kindergarten in 1999-2000 to be able to conduct an analysis for this grade/year.

period under study (1999-2000 through 2003-2004). The reasoning behind this decision was that FRPL status is a proxy for socioeconomic status, which encompasses a number of factors that may adversely impact student achievement. Further, although a student's status as receiving FRPL may change from one year to the next, this does not necessarily imply a significant change in socioeconomic status. Thus, to be conservative, we included students in the FRPL group if they had received FRPL status at any time from 1999-2000 through 2003-2004. This also resolved the problem of missing FRPL data for any one year, since a student's FRPL status for purposes of the present study was based on whether he or she had received FRPL for any year between 1999-2000 and 2004-2005.

## **5. Identify Pre-K Students in the Assessment Data**

The final step in the data management process was to identify which students in the assessment datasets attended Pre-K. To do so, the assessment datasets were merged together with the Pre-K demographic file. We then identified the year that students were recorded as having attended Pre-K (specifically, at age 4) and examined whether that information was consistent with the grade(s) and year(s) in which they had assessment data, allowing for the possibility of grade retention. In other words, students who had assessment scores as Kindergarteners in 1999-2000 would be expected to have attended Pre-K in 1998-1999. As another example, students who had assessment scores as first graders in 2001-2002 would be expected to have attended Pre-K in 1999-2000, or 1998-1999 for those who repeated Kindergarten or were delayed in starting Kindergarten. In the majority of cases, the grade(s) and year(s) for which Pre-K students had assessment scores were consistent with the year in which they were identified as having attended Pre-K. There were, however, some discrepancies. Specifically, there were a small number of cases where the year the student was recorded as having attended Pre-K was later than the grade and year in which they had assessment scores. For example, a student who supposedly attended Pre-K in 2002-2003 but had assessment scores as a first grader in 2000-2001, a scenario which is not possible. For a number of these instances, we compared the demographic information reported in the Pre-K demographic file (when available) with the assessment data and EIS data (again, when available) and in many cases found that the information did not match. This indicates that although the ESSN in the different data sources is the same, the information does not belong to the same student. These cases were flagged and excluded from analysis. It is important to note again, however, that these instances were fairly rare.

The resulting number of Pre-K students with assessment data available for analysis in each grade/year is relatively small, especially in the early grade levels and years. Table 4 displays the number of Pre-K students in the Pre-K Demographic File (limited to those with ESSNs) and the number of Pre-K students with assessment scores for each grade/year covered in this report. The total number of four-year olds listed in the Pre-K Demographic file is included in parentheses for each cohort.

The reader should keep in mind that the numbers of Pre-K students in Table 4 will be smaller than those presented in Table A1 in Appendix A, as Table A1 includes counts of all Pre-K students enrolled each year, including 3- and 4- year olds, and estimates for the 2007-2008 academic year. Table 4 reflects the number of valid records available for analysis at the conclusion of the data management phase of this analysis.

**Table 4. Number of Pre-K Students in the Pre-K Demographic File and the Number of Pre-K Students Available for Analysis in Each Grade/Year Covered in this Report**

Cohort 1: Pre-K in 1998-1999 (n=319)				
Grade K	Grade 1	Grade 2	Grade 3	Grade 4
14	57	114	142	141
Cohort 2: Pre-K in 1999-2000 (n=316)				
Grade K	Grade 1	Grade 2	Grade 3	Grade 4
56	158	241	251	--
Cohort 3: Pre-K in 2000-2001 (n=1,115)				
Grade K	Grade 1	Grade 2	Grade 3	Grade 4
67	360	434	--	--
Cohort 4: Pre-K in 2001-2002 (n=2,282)				
Grade K	Grade 1	Grade 2	Grade 3	Grade 4
184	779	--	--	--
Cohort 5: Pre-K in 2002-2003 (n=2,555)				
Grade K	Grade 1	Grade 2	Grade 3	Grade 4
257	--	--	--	--

The primary reason for the relatively small number of Pre-K students in each grade/year is a lack of assessment data, especially for Kindergarten and first graders. As discussed previously, only about 20% of Kindergarten students were assessed between 1999-2000 and 2003-2003 and about 50% in First Grade. We also must consider, however, whether Pre-K students may be more likely to be from LEAs that do not assess students in Grades K-2. As one example, of the 319 students in the Pre-K demographic file who are reported to have attended Pre-K (age 4) in 1998-1999, only 37 (12%) attended Kindergarten in an LEA that assessed Kindergarten students in 1999-2000. As another example, of the 316 students in the Pre-K demographic file who are reported to have attended Pre-K (age 4) in 1999-2000, only 53 (17%) attended in an LEA that assessed Kindergarten students in 2000-2001. It is clear that some number of students changed LEAs, and some number of students may have entered Kindergarten late or repeated a grade, placing them in a different cohort from which they started. Still, the point remains that Pre-K students appear to be less likely to attend school systems that administer assessments in Grades K-2.

Another reason for relatively small samples of Pre-K participants, as was discussed previously, is that some students whose records indicated gender or race discrepancies were excluded from analyses. Students were also excluded if they were found to have more than one set of scores in a particular school year at different grade levels. Receipt of special education services was another exclusion criterion (see the next section for more details). Finally, some natural attrition is also to be expected (i.e., students moving and leaving the state school system).

After all due care was taken to identify Pre-K participants and locate their assessment scores for each year, the final step in preparing the data for analysis was to select a comparative sample of students from the dataset for each grade/year that did not attend Tennessee's Pre-K program.

## ***Sampling Strategy***

In order to evaluate the short- and long-term impact of Pre-K on student outcomes, Pre-K students must be compared to a similar group of students that did not attend Tennessee's Pre-K program. In our Preliminary Report, SRG proposed obtaining the non-Pre-K groups from the 2005-2006 EIS, using proportionate probability sampling to ensure that each cohort sample of non-Pre-K students proportionately resembled each cohort of Pre-K attendees with regard to important student characteristics—namely, gender, race, FRPL status, and special education status. Increased familiarity with the data sources since the submission of the Preliminary Report, however, has led us to slightly revise our strategy in a way we feel offers several advantages over the original proposed approach.

Although the selection of the Pre-K and non-Pre-K groups will be discussed in more detail below, one general change in our sampling strategy is that, rather than selecting students from the 2005-2006 EIS data, they are selected from the assessment data. We had originally proposed, for example, selecting the sample of non-Pre-K students who were Kindergarteners in 1999-2000 from the population of Sixth Grade students in the 2005-2006 EIS data, the sample of non-Pre-K students who were Kindergarteners in 2000-2001 from the population of Fifth Grade students in the 2005-2006 EIS data, and so on. The main rationale for originally proposing to select students using the EIS data was because the EIS data also contain demographic information necessary for sampling. The assessment data files, however, also include the same student characteristics.

Given a choice between using the assessment data or the EIS data for sample selection, then, the assessment data offers clear advantages. First, it does not assume normal progress of students. Second, it allows for student attrition. Using the EIS data, on the other hand, would restrict the samples to those students who: 1) were not held back at any time before 2005-2006, and 2) did not leave the Tennessee school system at any time before 2005-2006. Further, using the EIS data for sampling could also result in selecting students who had not taken assessment tests in Tennessee during any previous school year included in this report—for example, students who had just moved to Tennessee in the 2005-2006 school year. Finally, not all students who have assessment data in any year between 1999-2000 and 2003-2004 are in the 2005-2006 EIS.

A second change in the sampling strategy is that instead of selecting non-Pre-K samples that mirror the Pre-K groups with regard to the distribution of gender, race, FRPL status, and special education, we selected the non-Pre-K samples such that they mirror the Pre-K groups with regard to school district, gender, race, and FRPL status.

After becoming more familiar with the data, we determined it was important to match the non-Pre-K sample to the Pre-K groups' school district distribution. Because the non-Pre-K samples of students are chosen from the assessment data, by definition they are similar to the students who attended Pre-K in that they attended a school district that opted to administer assessments for that particular grade/year. Because the decision to test in Kindergarten, First, and Second Grades is largely budget-driven (LEAs must absorb the costs of these assessments), school districts that administered assessments in Grades K, 1, and or 2 are likely to have larger budgets than those that did not. This is not a guarantee, however; for example, schools may

have different funding priorities regardless of budget size. There are likely to be other important differences across the school districts that administer K-2 assessments as well. Thus, choosing samples of non-Pre-K students that mirror Pre-K students with regard to school district offers a greater degree of assurance that the Pre-K and non-Pre-K students are similar in key ways aside from individual characteristics (i.e., gender, race, and FRPL status). This also helps to address some potential statistical concerns about the amount of missing data in the K-2 assessments.

Another sampling issue that was re-evaluated was the inclusion of students' special education status. Initially we considered examining the impact of Pre-K on students receiving special education services, but after careful consideration it was decided to exclude these students from subsequent analyses. The rationale for this decision is twofold. First, given the generally small number of students in this category (generally less than 10% in each grade/year), we were concerned about small sample sizes in terms of violation of statistical test assumptions and low statistical power. A second and more conceptual reason deals with the nature of the Pre-K program. This is not a program aimed specifically at students requiring special education services; rather the program goal is to improve school readiness for entry into Kindergarten. Therefore, it is unclear the extent to which participation in Pre-K impacts a student's need for special education services in Kindergarten and beyond. This particular question seems to require additional data beyond that which is available for the present evaluation, including more information about the students' characteristics and the services they receive. Further, lacking additional information about students receiving special education services, it is unclear whether a reasonable comparison group could be obtained (i.e., students receiving special education services who did not attend Pre-K). Therefore, to reduce possible sources of error in a comparison of outcomes between students who participated in Pre-K and at-risk students who did not participate in Pre-K, we ultimately decided to exclude from the analysis students who participated in Pre-K and were identified as requiring special education services in a particular year for which assessment data were available.

The sampling strategy for the non-Pre-K samples involved creating a distribution of the Pre-K group for each class/grade by district, then by FRPL status within each district, then by race and gender. The goal was to create a sample of non-Pre-K students that resembled the Pre-K students as closely as possible in terms of their school district, FRPL status, race, and gender. Of note, although great care was taken to make sure that the race (in aggregate) of the non-Pre-K group matched the race of the Pre-K group, in some instances there are slight differences due to small sample sizes for some ethnicities (American Indian/Native American, Asian/Pacific Islander, and Hispanic). However, for purposes of the analysis, we created two categories for race—white and non-white. Again, this was necessary because of the small number of students in some racial/ethnic groups. For all the samples, the proportion of white and non-white students was statistically equivalent. Within the category “non-white,” the non-Pre-K samples were constructed to match the composition of the Pre-K groups as closely as possible.<sup>8</sup>

At this point, as many Pre-K students as possible had been identified in the assessment data, and any inaccuracies or irregularities were resolved. Then, a comparable sample of non-Pre-K students was selected for each cohort and grade level in the timeframe under investigation. The

---

<sup>8</sup> Appendix D provides a descriptive overview of the demographic characteristics of the Pre-K groups (each cohort and grade level) included in this report. The reader will note that the number of students assessed each year changes substantially, which is consistent with the overall trend for assessments in Tennessee in grades K-2.

next step was to conduct the appropriate statistical analysis to determine whether there were meaningful differences, in aggregate, between the Pre-K and non-Pre-K groups.

## ***Analytic Approach***

For student assessment data between 1999-2000 and 2003-2004, separate analyses were conducted for each cohort at each grade level. As was discussed previously, many Pre-K students attended schools that did not administer assessments in Grades K-2 (particularly in Kindergarten). Further, some inaccuracies and inconsistencies in the data resulted in excluding some cases from the analysis. Thus, longitudinal analyses (i.e., analyses using data from the same students measured at multiple points in time) were simply not feasible for this report because of the small number of students who were assessed at least twice within this time period. We expect, however, that longitudinal analysis will be possible for subsequent reports, particularly for Grades Three through Five because testing is mandatory for those grades.

Analyses were restricted to students who had received FRPL at least one year in the time period under study (1999-2000 through 2003-2004).

For each grade level within each cohort, Analysis of Variance (ANOVA) was used to test for differences in assessment scores as a function of participation in Pre-K. Separate analyses were conducted for each assessment (reading, language, mathematics, etc.). ANOVA allowed us to examine whether scores varied as a function of participation in Pre-K (comparing Pre-K participants versus non-Pre-K participants), and also whether there were relative differences as a function of race and/or gender (i.e., an interaction). The results indicate the probability or likelihood that the differences observed between the groups are due to chance; a standard significance level of  $p > .05$  was applied. A statistically significant interaction indicates that the nature of the relationship between two or more variables changes depending on one (or more) of the other factors. That is, when a statistical interaction is present, it means that the pattern of results observed for one (or more) group(s) differs from that of another group. For example, an interaction between Pre-K participation and gender might indicate that Pre-K is associated with better assessment scores for students of one gender but not another.

In total, there were fourteen grade/years for which analyses were performed. Given the number of analyses performed, it would be prohibitive to discuss each one individually. Thus, the results section provides a summary of the statistically significant findings—or the group differences that are highly unlikely to be obtained as a result of chance or sampling error.

## **Results**

### **Cohort 1**

#### *Kindergarten*

Only a small number of students who participated in Pre-K in 1998-1999 were assessed in Kindergarten in the 1999-2000 school year. Given that there were only 14 valid records for the Pre-K group, no statistical analysis was possible for this particular grade level for Cohort 1.

#### *First Grade*

There were no statistically significant effects involving participation in Pre-K, and Pre-K participation did not differ by either race or gender among First Grade students in this cohort.

### Second Grade

There were no statistically significant effects involving participation in Pre-K, and Pre-K participation did not differ by either race or gender among Second Grade students in this cohort.

### Third Grade

For this grade level of this cohort, there were a number of statistically significant effects involving participation in the Pre-K program.

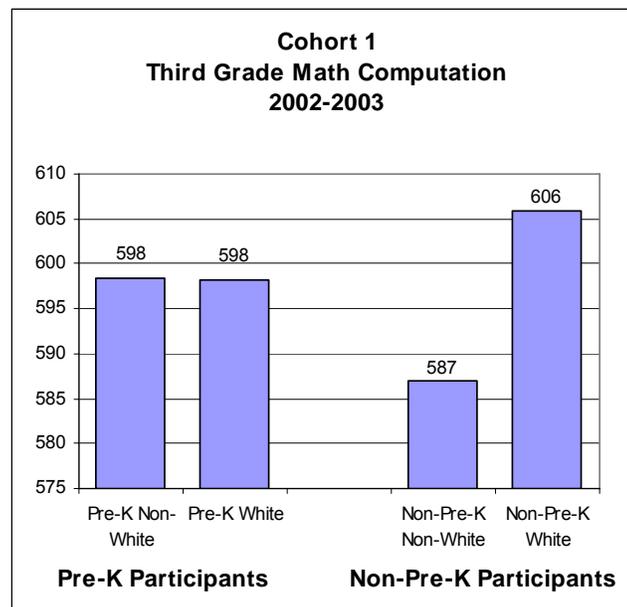
First, according to the results of the **science** assessment, students who participated in Pre-K scored higher, on average (mean = 621) compared with a similar group of students who did not participate in Pre-K (mean = 614).

There was a similar finding for the results of the **word analysis** assessment. Students who participated in Pre-K scored higher, on average (mean = 644) on word analysis compared with a similar group of students who did not participate in Pre-K (mean = 634).

Another significant effect was observed for **social studies**. Students who participated in Pre-K scored higher, on average (mean = 635) compared with a similar group of students who did not participate in Pre-K (mean = 625).

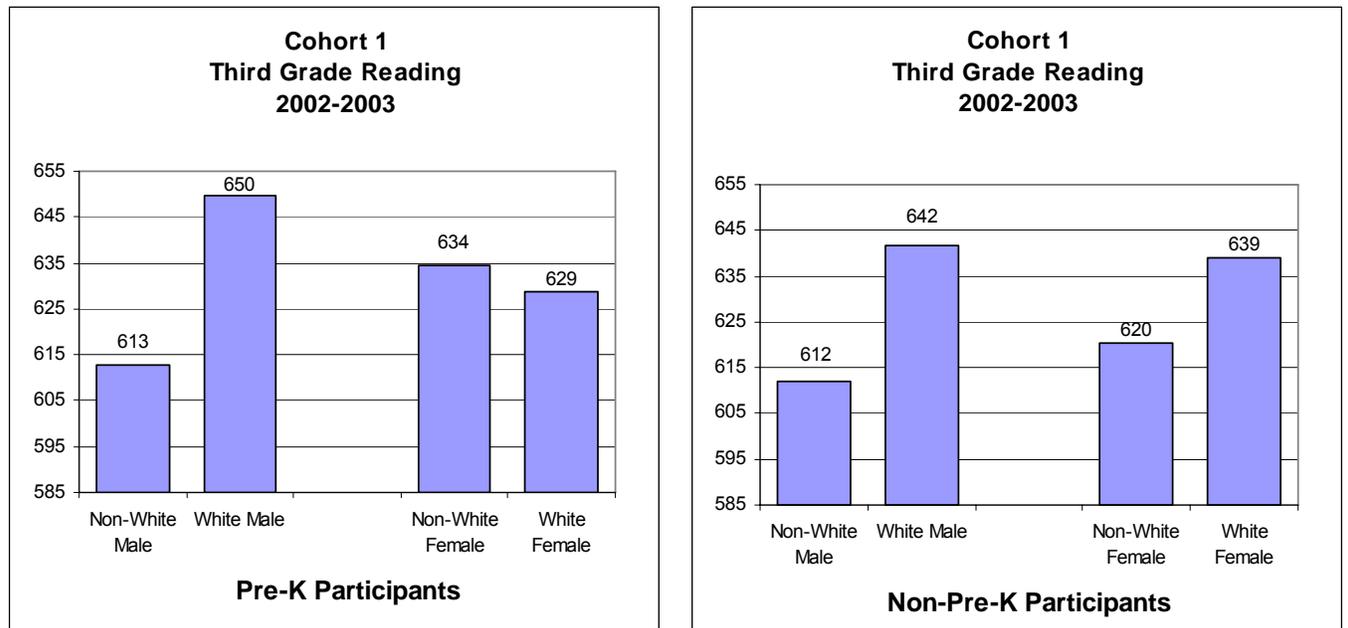
There were two additional and more complex effects involving participation in Pre-K in this analysis. First, the results of the **math computation** assessment revealed a benefit of the Pre-K program for non-white students. Specifically, non-white students who participated in Pre-K obtained a higher mean score on math computation than non-white students who did not participate in Tennessee's Pre-K program. Further, the scores of non-white students, on average, did not differ from white students who participated in Pre-K, unlike the significant difference between white and non-white students seen for the non-Pre-K students (see Figure 1, below).

**Figure 1. Third Grade Math Computation Mean Scores (2002-2003) by Pre-K Participation and Student Race**



Second, there was a benefit for non-white females reflected in the **reading** assessment results. Figure 2 shows the mean reading assessment scores by race and gender for both the Pre-K and non-Pre-K students. Although there were no statistically significant differences as a function of Pre-K participation among non-white males, white males, and white females, non-white female Pre-K participants scored higher in reading than non-white females who did not participate in Pre-K.

**Figure 2. Third Grade Reading Mean Scores (2002-2003) by Pre-K Participation, Student Gender, and Student Race**

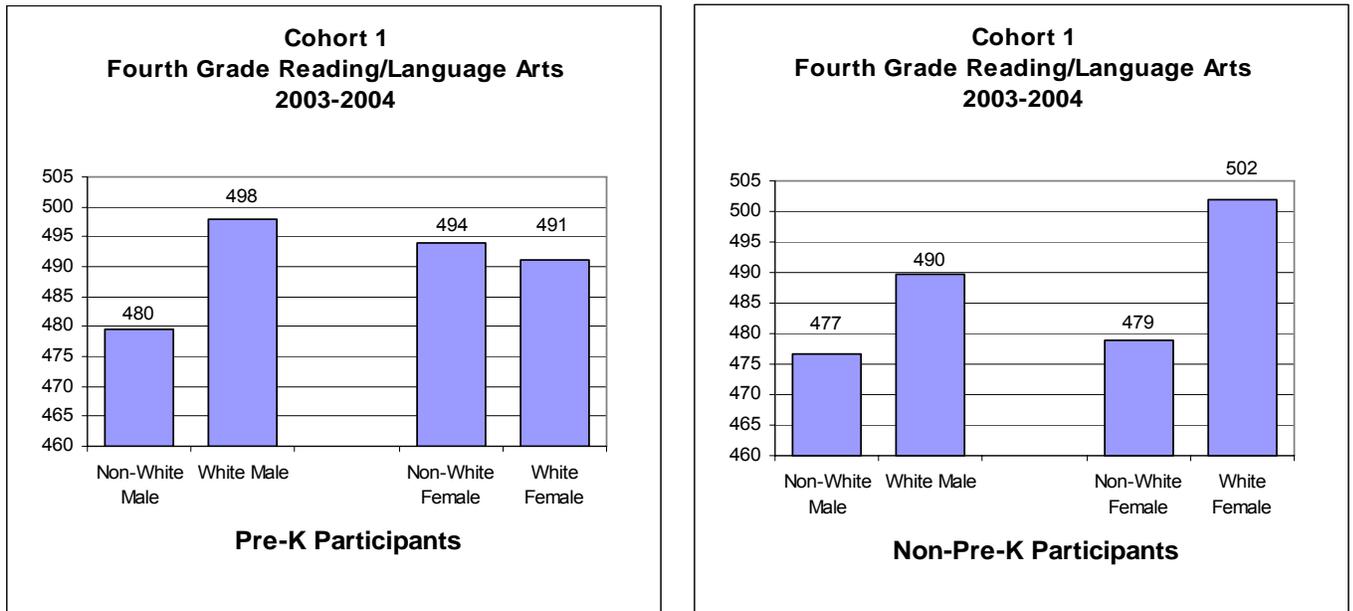


#### *Fourth Grade*

Cohort 1 students assessed in the Fourth Grade (2003-2004) took a Criterion-Referenced Test (CRT) for the first time. This is significant because the nature and interpretation of the test is different than for previous years' assessments. Fourth Grade students in Cohort 1 were assessed in **reading/language arts, mathematics, science, and social studies** using CRTs.

Although there were no differences associated with Pre-K participation in assessment scores for mathematics, science, and social studies, there was a significant effect for **reading/language arts**. Like the pattern obtained for reading/language arts scores among students in this cohort in the Third Grade, Fourth Grade reading scores also differed as a function of Pre-K participation, race, and gender. Specifically, non-white females who participated in Pre-K scored higher, on average, compared with non-white females who did not participate in Pre-K. Among students who did not participate in Pre-K, non-white females had a significantly lower mean reading/language arts score than white female students. Among Pre-K participants, however, there was not a statistically significant difference between non-white females and white females on the Fourth Grade reading/language arts assessment (see Figure 3).

**Figure 3. Fourth Grade Reading/Language Arts Mean Scores (2003-2004) by Pre-K Participation, Student Gender, and Student Race**



## Cohort 2

### *Kindergarten*

There were no statistically significant effects involving participation in Pre-K, and Pre-K participation did not differ by either race or gender among First Grade students in this cohort.

### *First Grade*

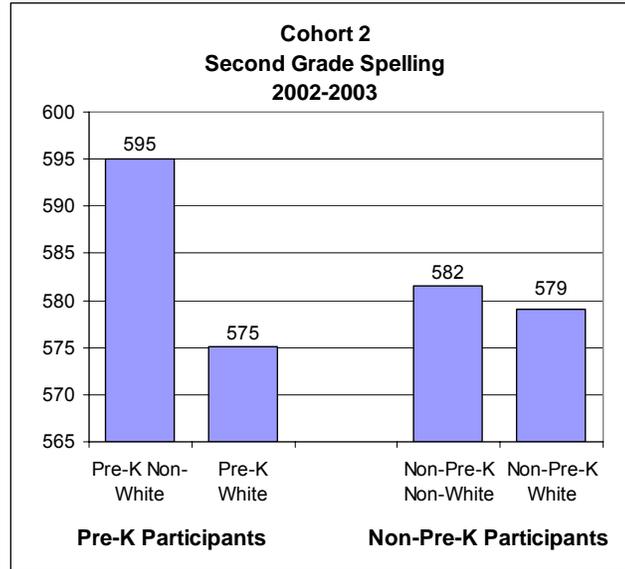
There was one statistically significant effect involving Pre-K participation. Students who participated in Pre-K tended to score higher on the **mathematics** assessment (mean = 547) than at-risk students who did not participate in Pre-K (mean = 539). There were no other significant effects of Pre-K participation for Cohort 2 students tested in the First Grade.

### *Second Grade*

When Cohort 2 students were tested in Second Grade, there were a number of statistically significant effects as a function of Pre-K participation. Pre-K participation showed beneficial results for **reading, language, and language mechanics** assessments. For reading, Pre-K participants scored higher (mean = 619) compared with non-Pre-K participants (mean = 612). Similarly, on the language assessment, Pre-K participants scored higher (mean = 625) than non-Pre-K participants (mean = 617). The same pattern was observed for language mechanics: Pre-K students scored higher (mean = 620) compared with their non-Pre-K peers (mean = 613).

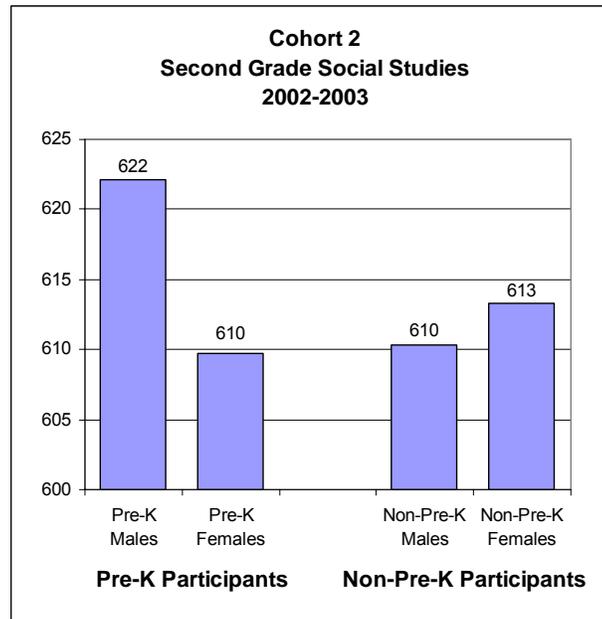
Results for the **spelling** assessment differed for Pre-K and non-Pre-K students with respect to race in this cohort in Second Grade. Non-white students who participated in Pre-K scored higher on the spelling assessment compared with non-white students who did not participate in Pre-K. Among students who participated in Pre-K, non-white students scored higher on spelling than white Pre-K participants whereas there was no significant difference seen among the non-Pre-K students (see Figure 4).

**Figure 4. Second Grade Spelling Mean Scores (2002-2003) by Pre-K Participation and Student Race**



The last effect found among students in Cohort 2, Grade 2 was a significant difference between Pre-K and non-Pre-K students based on gender for the **social studies** assessment. Male students who participated in Pre-K scored higher than male students who did not participate in Pre-K. Results also showed that among Pre-K participants, male students scored higher, on average, than females— whereas no such pattern was seen for the non-Pre-K students (see Figure 5).

**Figure 5. Second Grade Social Studies Mean Scores (2002-2003) by Pre-K Participation and Student Gender**



### Third Grade

When Cohort 2 students were assessed in the Third Grade (2003-2004), they began taking Criterion-Referenced tests (CRTs) for **reading/language arts**, **mathematics**, **science**, and **social studies**.

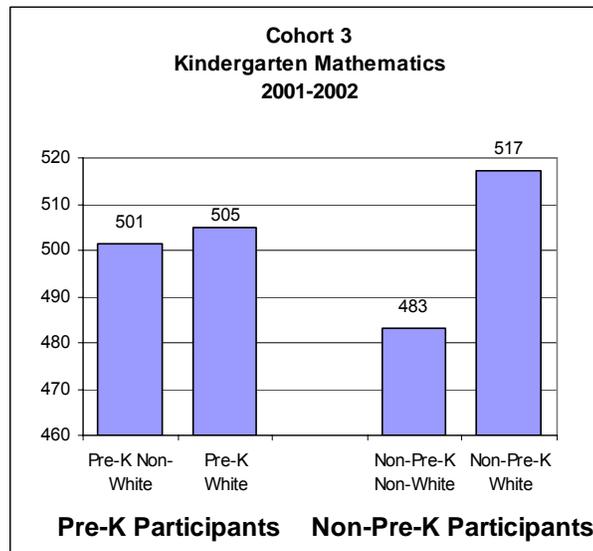
The analyses of these four assessments for Cohort 2 Third Grade students revealed only one significant effect of Pre-K participation. On the **social studies** assessment, students who participated in Pre-K tended to score slightly higher (mean = 198) compared with students who did not participate in Pre-K (mean = 195).

### Cohort 3

#### Kindergarten

There was one significant effect of Pre-K participation among students in Cohort 3 when they were assessed in Kindergarten. Pre-K participation showed a benefit for non-white students for the **mathematics** assessment. Non-white students who participated in Pre-K scored significantly higher, on average, compared with non-white students who did not participate in Pre-K. Also, as Figure 6 shows, among Kindergarten students who did not participate in Pre-K, white students scored higher than non-white students. Among students who participated in Pre-K, however, there was no statistically significant difference between white and non-white students.

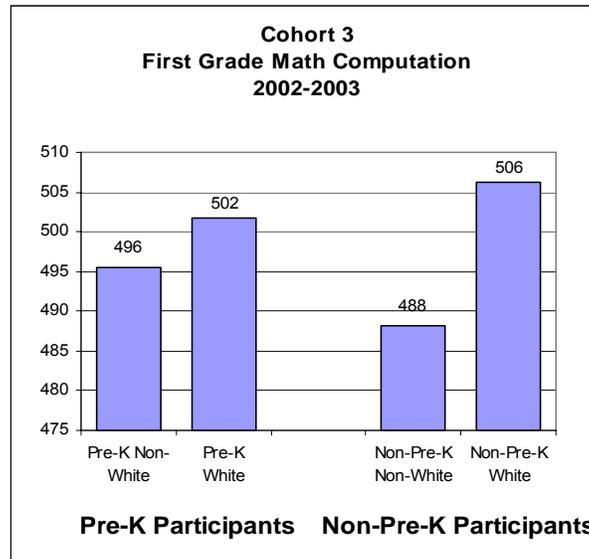
**Figure 6. Kindergarten Mathematics Mean Scores (2001-2002) by Pre-K Participation and Student Race**



### First Grade

When Cohort 3 students were assessed in First Grade, the results for race were different for Pre-K and non-Pre-K students for the **math computation** assessment (see Figure 7). Similar to the results obtained in Kindergarten, among the non-Pre-K participants, white students scored higher than non-white students. There was no statistically significant difference between the mean scores for white and non-white students among the Pre-K participants, however.

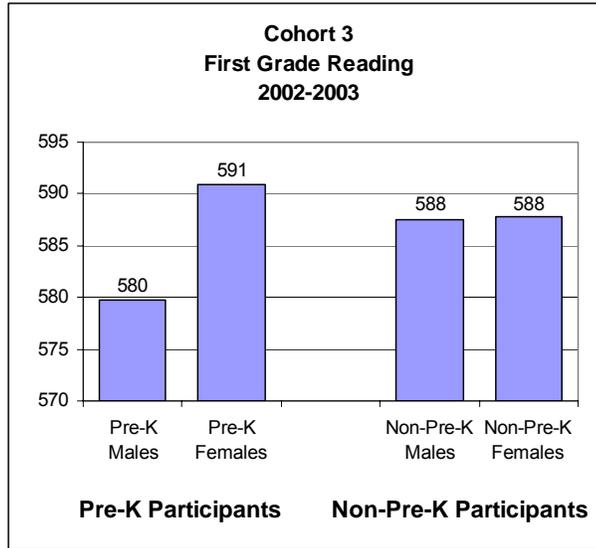
**Figure 7. First Grade Math Computation Mean Scores (2002-2003) by Pre-K Participation and Student Race**



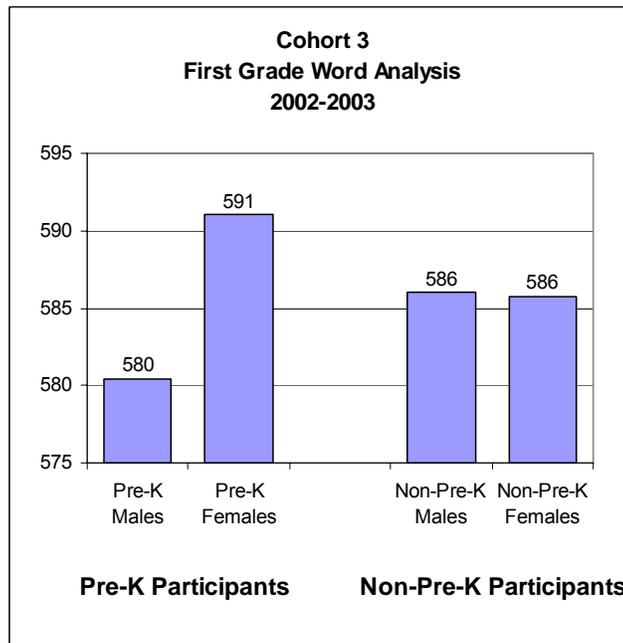
For six of the First Grade assessments, analyses revealed a difference between Pre-K and non-Pre-K students based on gender. These assessments include **reading, word analysis, language, mathematics, math computation, and social studies**.

The analysis for **reading, word analysis, and language** showed the same general pattern of results. Although there were no significant differences between Pre-K and non-Pre-K male students or female students for any of these three tests, among Pre-K participants, there was a significant difference between male students and female students such that Pre-K females scored higher on these assessments. This gender difference was not present for the non-Pre-K group (see Figures 8, 9, and 10).

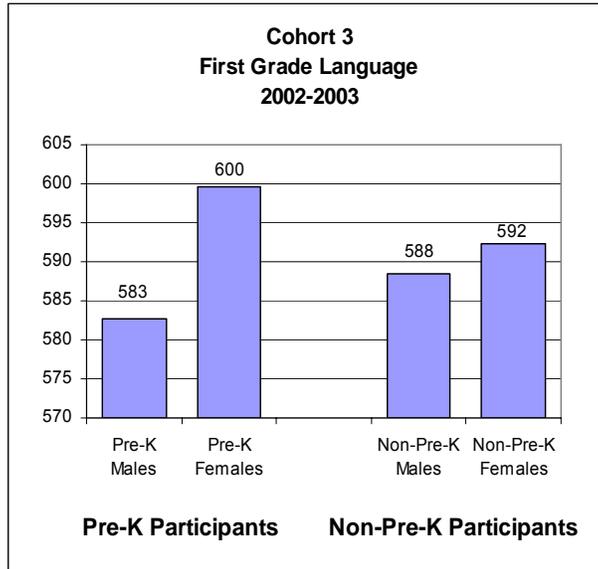
**Figure 8. First Grade Reading Mean Scores (2002-2003) by Pre-K Participation and Student Gender**



**Figure 9. First Grade Word Analysis Mean Scores (2002-2003) by Pre-K Participation and Student Gender**

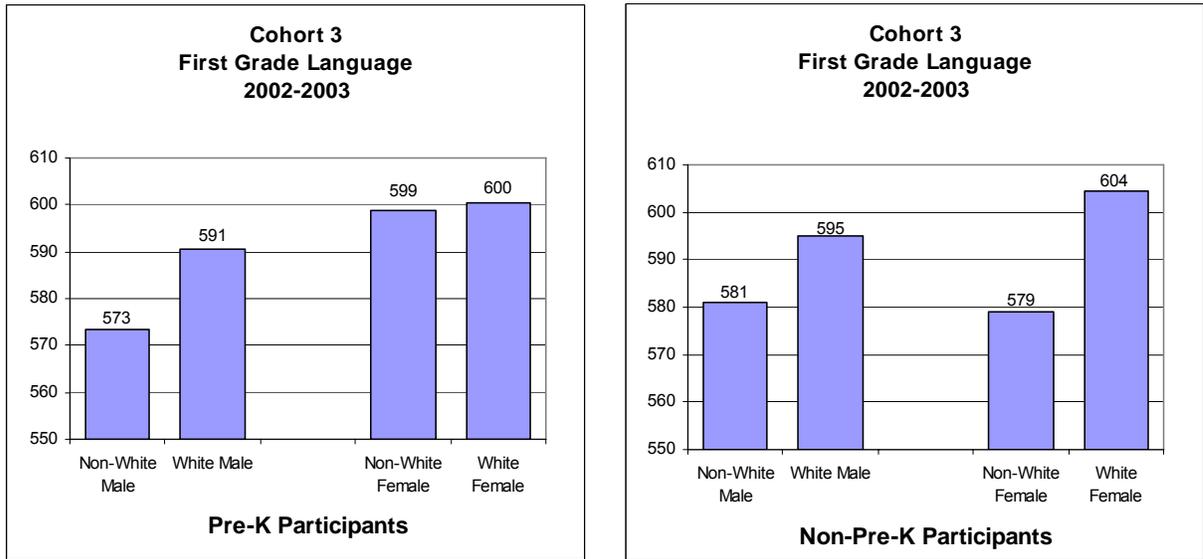


**Figure 10. First Grade Language Mean Scores (2002-2003) by Pre-K Participation and Student Gender**



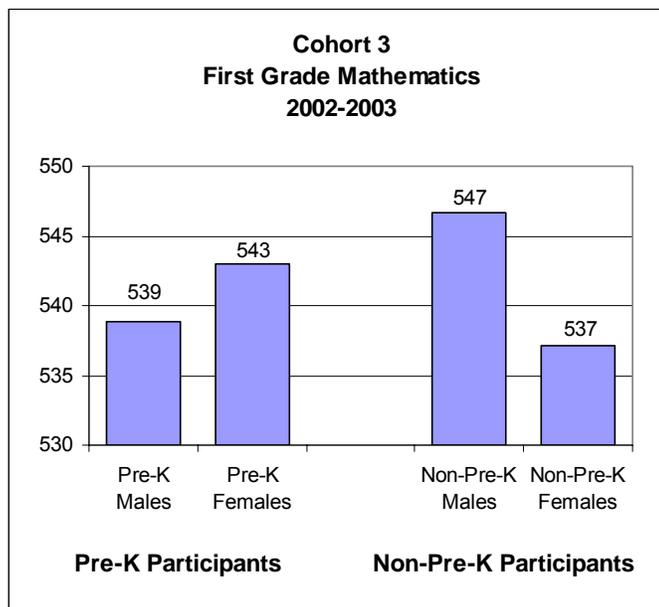
For the **language** assessment the difference in the gender results (Figure 10) was further qualified by a dependence on race. That is, the pattern of results changes when student race is taken into consideration. Non-white females who attended Pre-K scored higher on the language assessment than non-white females who did not attend Pre-K. Further, although white students tended to score higher than non-white students on the language assessment in both the Pre-K and the non-Pre-K groups, for females who participated in Pre-K there was no statistically significant difference between non-white and white students. That is, non-white females who participated in Pre-K tended to score, on average, as well as white females who participated in Pre-K. Comparatively, among at-risk female students who did not participate in Pre-K, white females tended to score higher than non-white females. This pattern of results is shown in Figure 11.

**Figure 11. First Grade Language Mean Scores (2002-2003) by Pre-K Participation, Student Race, and Student Gender**

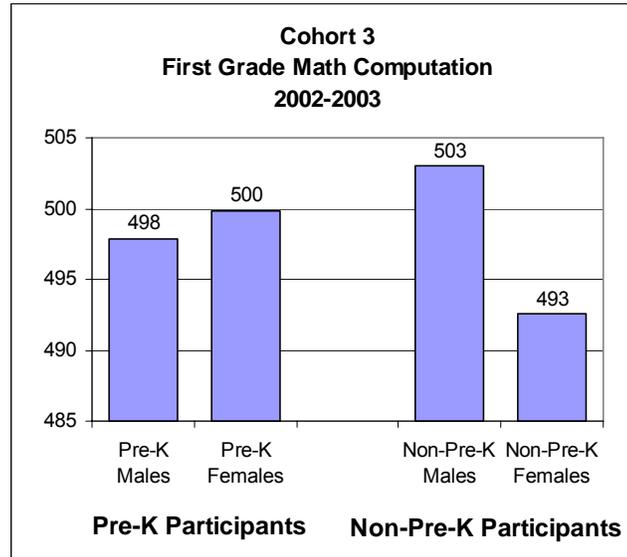


A slightly different pattern of results was observed for **mathematics** and **math computation**. For both of these assessments, the pattern of results by gender for Pre-K students differed from that of non-Pre-K students. Among non-Pre-K participants, male students tended to outscore female students. However, among Pre-K participants, there was not a statistically significant difference between male and female student scores, on average. For mathematics, male Pre-K students tended to perform slightly lower than their non-Pre-K counterparts, but for math computation there was not a significant difference between these groups (see Figures 12 and 13).

**Figure 12. First Grade Mathematics Mean Scores (2002-2003) by Pre-K Participation and Student Gender**

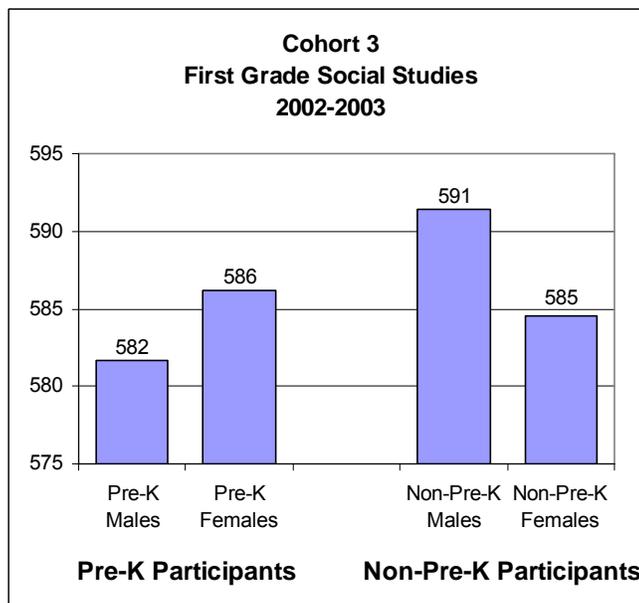


**Figure 13. First Grade Math Computation Mean Scores (2002-2003) by Pre-K Participation and Student Gender**



A pattern similar to that obtained for mathematics was also seen in the results of the analysis of **social studies** assessment results. Again, the gender pattern of results was different for the Pre-K group and the non-Pre-K group for social studies. Among non-Pre-K participants, male students tended to outscore female students. Among Pre-K participants, there was not a statistically significant difference between male and female student scores. However, looking between groups, male Pre-K students performed significantly lower on the social studies assessment than their non-Pre-K counterparts (see Figure 14).

**Figure 14. First Grade Social Studies Mean Scores (2002-2003) by Pre-K Participation and Student Gender**



### Second Grade

When Cohort 3 students were tested in the Second Grade, there were no statistically significant effects involving participation in Pre-K, and Pre-K participation did not differ by either race or gender among Second Grade students in this cohort.

## Cohort 4

### Kindergarten

There were no statistically significant effects involving participation in Pre-K, and Pre-K participation did not differ by either race or gender among First Grade students in this cohort.

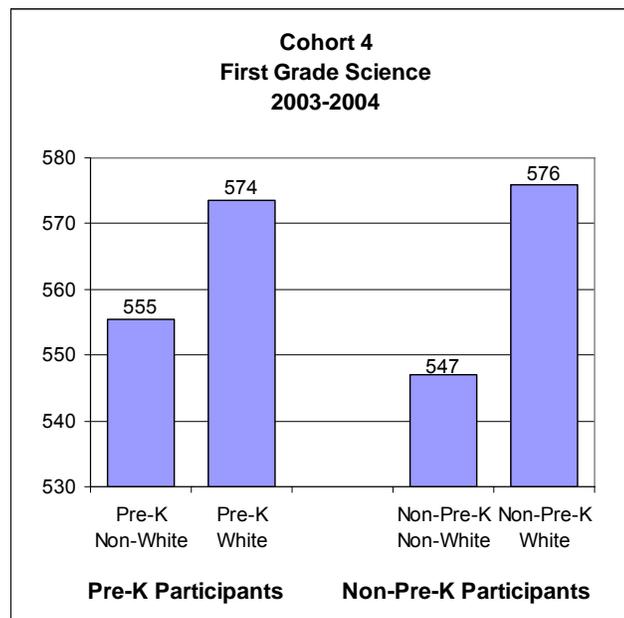
### First Grade

There were five assessments that varied as a function of Pre-K participation: **social studies, science, reading, vocabulary, and word analysis.**

First, in terms of **social studies**, Pre-K participation was found to be beneficial. Pre-K students in Cohort 4, First Grade scored higher in social studies (mean = 592) compared to their non-Pre-K counterparts (mean = 587).

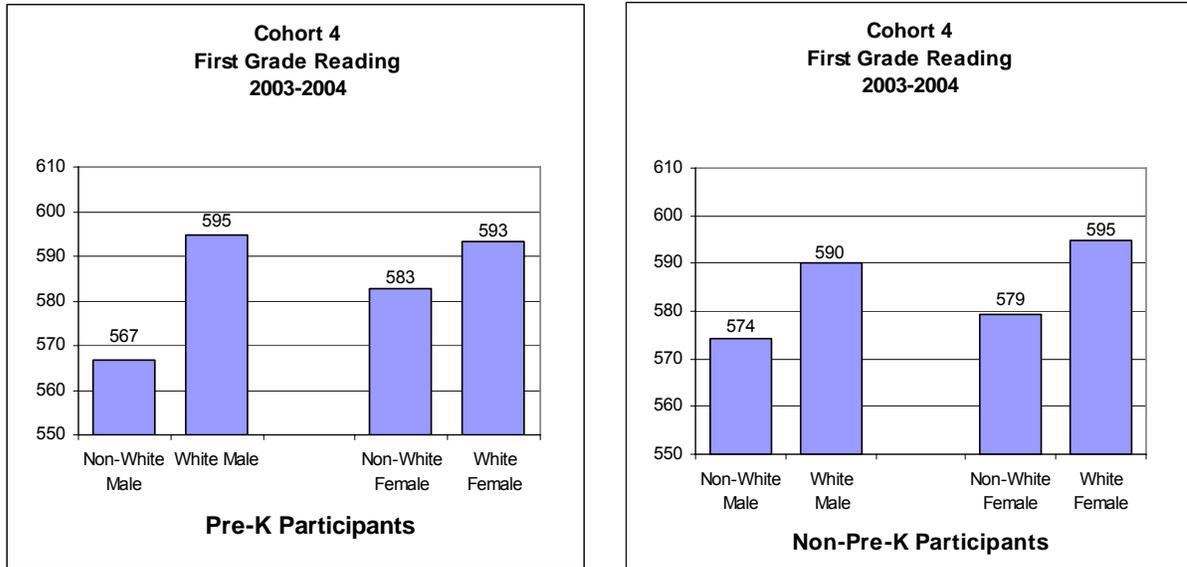
For **science**, there was a significant difference between non-white students who attended Pre-K and non-white students who did not attend Pre-K. Overall, white students tended to score higher on the science assessment than non-white students for both the Pre-K and non-Pre-K groups; however, non-white students who attended Pre-K scored higher than non-white students who did not attend Pre-K (see Figure 15).

**Figure 15. First Grade Science Mean Scores (2002-2003) by Pre-K Participation and Student Race**



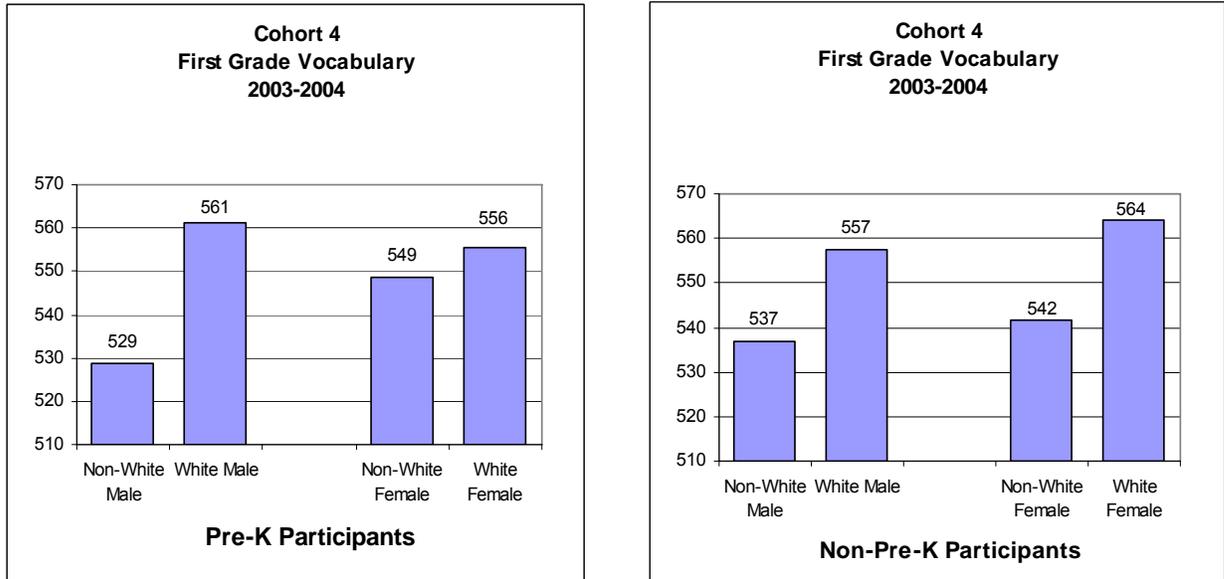
The analysis for **reading** revealed that the Pre-K and non-Pre-K groups had different patterns of results based on race and gender. These results are shown in Figure 16. In the case of this assessment, non-white males who participated in Pre-K actually scored lower, on average, compared with non-white males who did not participate in Pre-K. Further, there was a significant difference between white and non-white students, among both males and females, such that white students tended to score higher than nonwhite students.

**Figure 16. First Grade Reading Mean Scores (2003-2004) by Pre-K Participation, Student Race, and Student Gender**



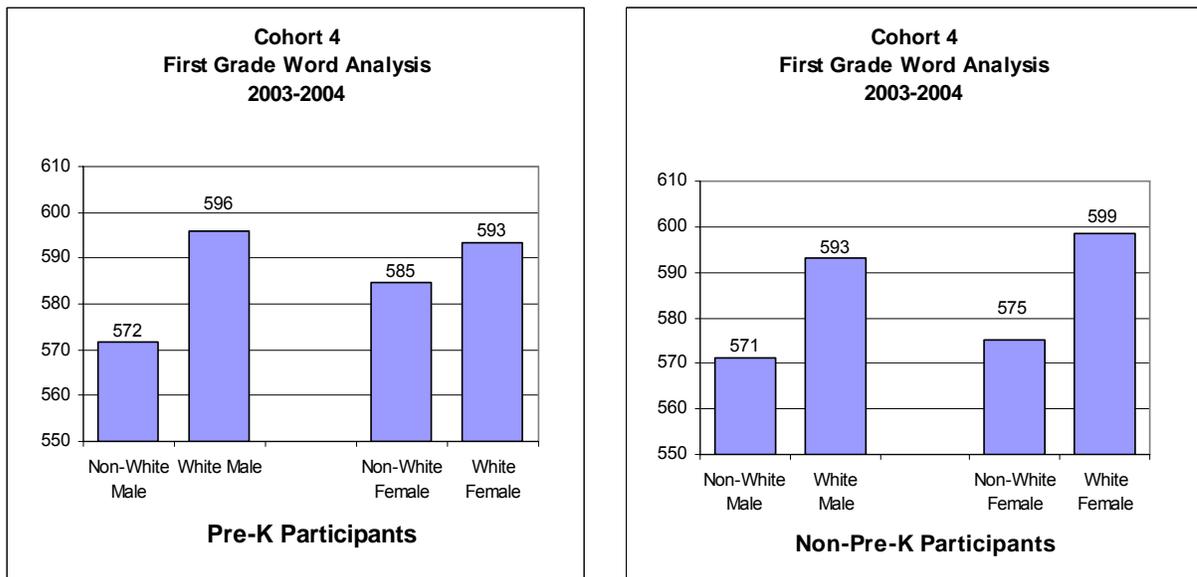
However, a different pattern of results—the same trend observed for earlier cohorts in reading and language assessments— was observed in analysis of the **vocabulary** assessment (see Figure 17). When Cohort 4 students were assessed in the First Grade, non-white females who attended Pre-K scored higher on the vocabulary assessment compared with non-white females who did not participate in Pre-K. Further, among Pre-K participants, non-white females scored at about the same level as white females on the vocabulary assessment, whereas for the non-Pre-K group non-white females scored lower than white females. This is the same pattern of results observed among Cohort 1 students in the Third Grade (reading) and Fourth Grade (reading/language arts), as well as for Cohort 3 students in the First Grade (language).

**Figure 17. First Grade Vocabulary Reading Mean Scores (2003-2004) by Pre-K Participation, Student Race, and Student Gender**



This same pattern of results was also found when analyzing the **word analysis** assessment for First Grade students in Cohort 4. That is, non-white females who participated in Pre-K scored higher in word analysis, on average, than non-white females who did not participate in Pre-K. Further, among Pre-K participants, non-white females scored at about the same level as white females on the word analysis assessment whereas for the non-Pre-K group non-white females scored lower than white females (see Figure 18).

**Figure 18. First Grade Word Analysis Mean Scores (2003-2004) by Pre-K Participation, Student Race, and Student Gender**

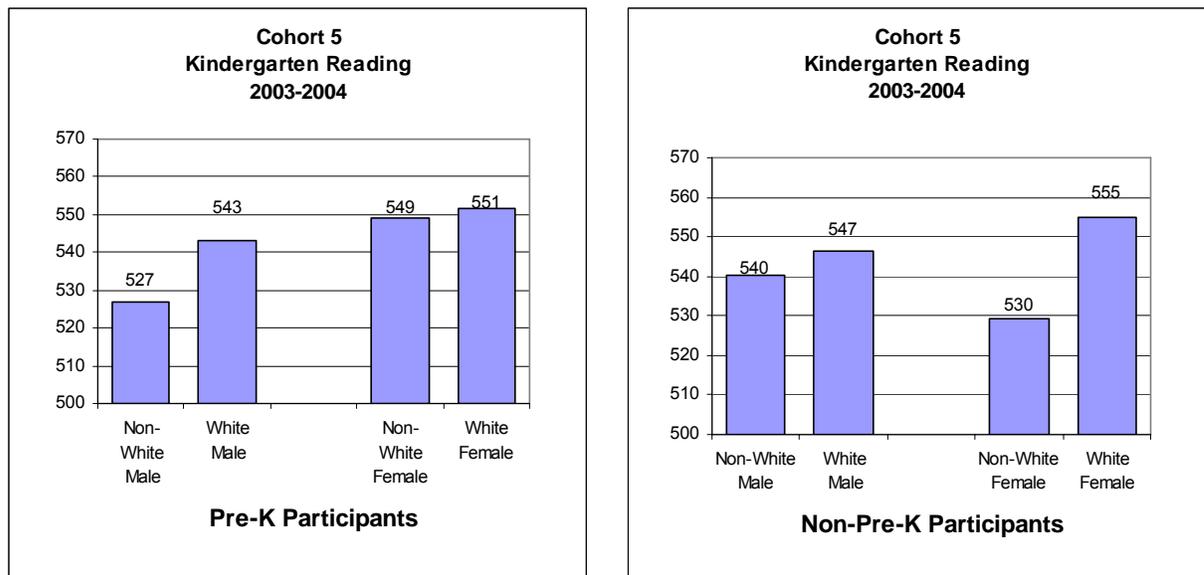


## Cohort 5

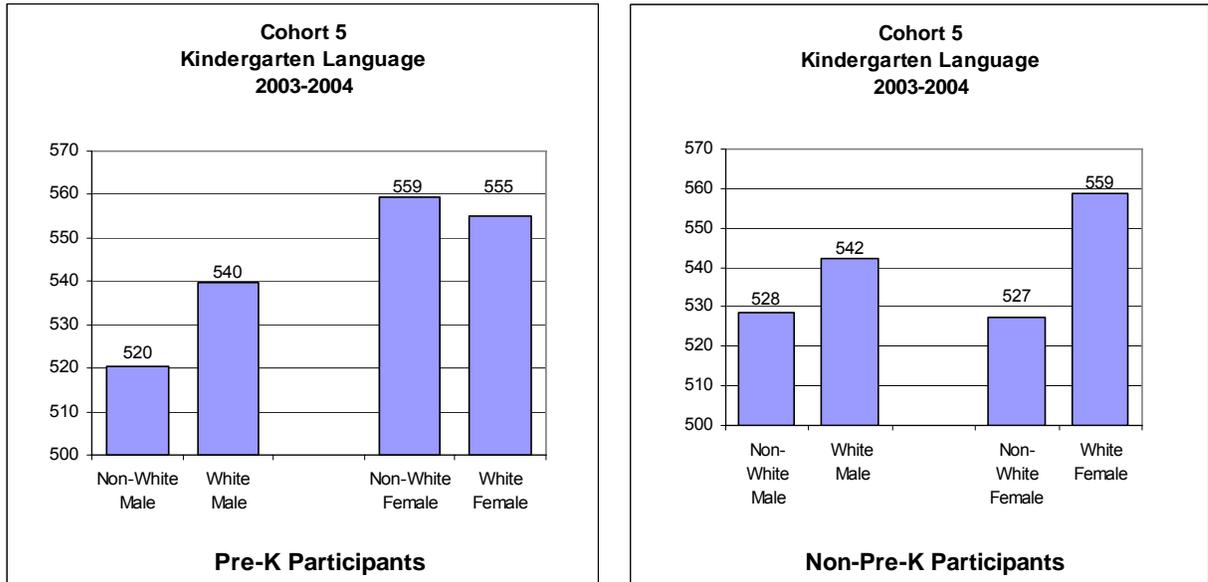
### Kindergarten

In analysis of Cohort 5 Kindergarten assessments, significant effects associated with Pre-K participation were found for **reading** and **language**. In the case of both assessments, the pattern of results based on gender differed between the Pre-K group and the non-Pre-K group, such that female Pre-K participants scored higher on reading and language compared with at-risk females who did not participate in Pre-K. Also in both cases, the results differed as a combination of race and gender. The same pattern observed in earlier analyses emerged in this group as well: Among Pre-K participants, non-white female students scored as well as white female students. Further, non-white females who participated in Pre-K scored higher than non-white females who did not participate in Pre-K. This recurrent pattern of results is shown in Figures 19 and 20.

**Figure 19. Kindergarten Reading Mean Scores (2003-2004) by Pre-K Participation, Student Race, and Student Gender**



**Figure 20. Kindergarten Language Mean Scores (2003-2004) by Pre-K Participation, Student Race, and Student Gender**



No other results indicated a statistically significant difference in scores associated with Pre-K participation.

## Discussion and Next Steps

Across analyses of these five cohorts of students over five years, there were numerous positive effects associated with Pre-K participation such that **there do seem to be some short-term and long-term benefits of Pre-K participation on assessments in Grades K-4**. In some cases, Pre-K participation is associated with higher scores on assessments. In other cases, only some groups of students seem to score higher on assessments after Pre-K participation, and this varies as a function of gender and race. The significant effects found in the course of this analysis are summarized in Table 4.

**Table 4. Summary of Significant Findings Involving Pre-K Participation in the First Interim Report**

Cohort 1	
Kindergarten (1999-2000)	<i>No analyses possible</i>
Grade 1 (2000-2001)	No effects of Pre-K participation
Grade 2 (2001-2002)	No effects of Pre-K participation
Grade 3 (2002-2003)	<ul style="list-style-type: none"> <li>Pre-K students scored higher, overall, on <b>Science, Social Studies</b> and <b>Word Analysis</b>, than other at-risk students who did not attend Pre-K.</li> <li>Non-white students who participated in Pre-K scored higher on <b>Math Computation</b> than non-white students who did not attend Pre-K.</li> <li>Non-white female students who participated in Pre-K scored higher on <b>Reading</b> compared to non-white female students who did not attend Pre-K, and scored as well as white female students who participated in Pre-K.</li> </ul>
Grade 4 (2003-2004) (note: CRT)	<ul style="list-style-type: none"> <li>Non-white female students who participated in Pre-K scored higher on <b>Reading/Language Arts</b> compared to non-white female students who did not attend Pre-K, and scored as well as white female students who participated in Pre-K.</li> </ul>
Cohort 2	
Kindergarten (2000-2001)	No effects of Pre-K participation
Grade 1 (2001-2002)	<ul style="list-style-type: none"> <li>Pre-K students scored slightly higher, overall, on <b>Mathematics</b> than other at-risk students who did not attend Pre-K.</li> </ul>
Grade 2 (2002-2003)	<ul style="list-style-type: none"> <li>Pre-K students scored higher, overall, on <b>Reading, Language and Language Mechanics</b>, than other at-risk students.</li> <li>Non-white students who participated in Pre-K scored higher on <b>Spelling</b> compared with non-white students who did not attend Pre-K.</li> <li>Male students who participated in Pre-K scored higher on <b>Social Studies</b> than male students who did not attend Pre-K.</li> </ul>
Grade 3 (2003-2004) (note: CRT)	<ul style="list-style-type: none"> <li>Pre-K students scored slightly higher, overall, on <b>Social Studies</b> than other at-risk students who did not attend Pre-K.</li> </ul>

Table 4, cont'd

Cohort 3	
Kindergarten (2001-2002)	<ul style="list-style-type: none"> <li>Non-white students who participated in Pre-K scored significantly higher on <b>Mathematics</b> compared with non-white students who did not attend Pre-K.</li> </ul>
Grade 1 (2002-2003)	<ul style="list-style-type: none"> <li>Non-white students who participated in Pre-K scored significantly higher on <b>Math Computation</b> compared with non-white students who did not attend Pre-K.</li> <li>On <b>Mathematics</b> and <b>Math Computation</b>, male students who attended Pre-K scored <u>lower</u> than male non-Pre-K participants</li> <li>Among Pre-K participants, females scored higher than males on <b>Reading, Word Analysis, and Language</b>.</li> <li>For <b>Language</b>, non-white female students who participated in Pre-K scored higher compared to non-white female students who did not attend Pre-K, and scored as well as white female students who participated in Pre-K.</li> <li>Male Pre-K students performed significantly <u>lower</u> on <b>Social Studies</b> than at-risk males who did not attend Pre-K.</li> </ul>
Grade 2 (2003-2004)	No effects of Pre-K participation
Cohort 4	
Kindergarten (2002-2003)	No effects of Pre-K participation
Grade 1 (2003-2004)	<ul style="list-style-type: none"> <li>Pre-K students scored higher, overall, on <b>Social Studies</b> than other at-risk students who did not attend Pre-K.</li> <li>Non-white students who attended Pre-K scored higher on <b>Science</b> than non-white students who did not attend Pre-K.</li> <li>Non-white males who participated in Pre-K scored <u>lower</u> on <b>Reading</b> compared with non-white males who did not attend Pre-K</li> <li>For <b>Vocabulary and Word Analysis</b>, non-white females who attended Pre-K scored higher compared with non-white females who did not attend Pre-K, and scored as well as white female students who participated in Pre-K.</li> </ul>
Cohort 5	
Kindergarten (2003-2004)	<ul style="list-style-type: none"> <li>Among Pre-K participants, non-white female students scored as well as white female students on <b>Reading and Language</b>; non-white females who participated in Pre-K scored higher on these assessments than non-white females who did not attend Pre-K.</li> </ul>

The pattern of results generally seems to change between cohorts and grade levels, although there are some recurring patterns. For some cohorts in some years, Pre-K students score better than the comparison group of at-risk students who did not participate in Pre-K. The effects are seen in many assessments, including reading, language, word analysis, spelling, social studies, and science—although the benefits do not appear to be consistent for most of the assessments, and do not regularly seem to manifest themselves across multiple grade levels. There is one notable exception, though, and this consistent pattern is seen in reading and language arts.

For assessments of reading and language arts (both NRT and CRT assessments) there is a recurrent pattern of results such that **non-white females who attended Pre-K scored higher than non-white females who did not attend Pre-K**. Further, among students who participated in Pre-K, non-white female students scored comparably to white female students. This pattern

was not observed for non-white males, for white males, or for white females. This is an intriguing finding because it suggests that the non-white female students, in the aggregate, may reap the greatest benefits of Pre-K participation, particularly in the area of reading and language arts. Further, this pattern of results emerged over multiple cohorts (Cohorts 1, 3, 4, and 5), over multiple assessments (reading, language, vocabulary, and word analysis), and in multiple grades (K, 1, 3, and 4). This suggests that Pre-K participation may help to close the “achievement gap” that is often observed between white and non-white students—but the impact is localized such that non-white females, on average, tend to show the greatest benefits of Pre-K participation, at least with regard to certain assessments. It is both interesting and potentially important that minority students seem to benefit the most in the short- and long-term—particularly female students, and particularly in the area of reading/language arts. Further exploration of this potentially important finding will be necessary as the evaluation progresses.

For assessments involving mathematics, the results are less consistent but do reveal a positive impact of Pre-K participation, depending on grade level. For example, among Cohort 1 students in the Third Grade, Pre-K participation was associated with improved mathematics performance among non-white students relative to a comparison group of at-risk non-white students who did not participate in Pre-K—but this effect was not observed at the next grade level. Among Cohort 2 students, there was a significant difference in mathematics scores in First Grade such that Pre-K students scored higher on this assessment compared with non-Pre-K students, but again, this pattern was not evident in other grade levels for this cohort. A more complex pattern was observed for Cohort 3: although Pre-K participation appears to be associated with higher mathematics performance in Kindergarten and First Grade for non-white students, it is also associated with lower performance for male students. Considering these results together, **the impact of Pre-K participation on mathematics performance is variable and suggests a limited short-term impact** based on the analyses conducted to date.

Likewise, the benefits of Pre-K participation are limited in terms of science and social studies—these effects vary by cohort and grade level. There is some evidence to suggest that Pre-K participation is associated with higher performance on the social studies assessment, but in some cases Pre-K participation is associated with lower scores for social studies (Cohort 3, Grade 1).

## Considerations

Of course, interpretation of these results must be made with caution. Although there do seem to be benefits associated with participation in Pre-K and it may be tempting to attribute causality to participation in Pre-K to student outcomes in the short- and even in the long-term, there are many other important sources of variation that likely contribute to student assessment results that cannot be accounted for in the present study. Foremost, students were not randomly selected or assigned to attend Pre-K—parents selected this program for their children. Therefore, parental involvement is an important variable that unfortunately cannot be controlled for in the present study. Similarly, for the timeframe under investigation no classroom-level variables are available for inclusion in the analysis. Of course, it would be ideal to include as many potential sources of variation as possible as control variables, but unfortunately—at least for the First Interim Report—additional control variables are not available.

Although great care has been taken in the present analysis to utilize as much data as possible from Pre-K participants and to create valid comparison groups of at-risk students who did not attend Pre-K, there are still some factors that remain outside our control. First, although we can

be confident that the at-risk students in the comparison group did not attend Tennessee's Pre-K program, we cannot say conclusively that these students did not attend any Pre-K or early childhood enrichment program, and therefore we cannot control for that potential source of variation in the data.

A second challenge faced in the preparation of this report was small sample sizes for Pre-K participants, particularly in Grades K-2. Upon examination of data available, it appears that overall LEAs in Tennessee are unlikely to administer tests in Kindergarten and First Grade, and Pre-K participants may be particularly likely to attend LEAs that do not administer tests in these grades. Therefore, especially for the early years of the Pre-K program, it is difficult to gauge the true impact of Pre-K participation on student outcomes. Further, the fact that each grade level (K-2) includes different (increasing) numbers of Pre-K participants also may contribute to the lack of unifying findings across cohorts and grade level. Each analysis for cohort and grade level included more and more Pre-K students, finally leveling off in the Third Grade, when all students must complete these assessments. Examining the longer-term impact of Pre-K participation (Grades 3-5) may yield a clearer picture in terms of the impact of this early childhood enrichment given that all students must be assessed in Grades 3 and beyond,

## **Next Steps**

Further analyses are planned and this evaluation is ongoing. Many of the preliminary findings in this report must be explored further, and more extensive analyses become possible as more data become available.

The Second Interim Report (planned for March 2008) will compare student achievement between Pre-K and non-Pre-K students between the years 2004-2005 through 2006-2007. Over time, we do not anticipate to face the challenges experienced for this report in terms of small sample sizes. This is because of an increasing number of Pre-K participants each year, and more valid assessment data—at higher grade levels (Third, Fourth, and Fifth Grades) all students are required to complete standardized tests.

In the data to be analyzed for the Second Interim Report, five of the eight cohorts to be studied will be in the Third, Fourth, or Fifth Grades. We anticipate that for these grades, longitudinal analysis may be possible given that assessments in Grades Three through Five are mandatory and thus there will be larger numbers of students who can be followed over time. Additionally, because EIS data are available for two of the three years covered in the Second Interim Report, we expect to compare Pre-K and non-Pre-K students with regard to other indicators of academic achievement, such as attendance and disciplinary issues. The inclusion of multiple outcomes will offer a more thorough examination of the effectiveness of Tennessee's Pre-K program that was not possible for this First Interim Report.

Finally, the Second Interim Report will include an investigation of the relative impact of program characteristics on student outcomes. Ultimately, the analyses we conduct in this regard depend upon the availability of program-level data. However, one primary program characteristic that may be explored is a comparison of Pilot Pre-K programs and General Fund Pre-K sites. The two systems are alike in their classroom requirements regarding teacher credentials, class size, and curricular focus; however, they differ in their funding amount and source and their affiliation with the public school system (see Appendix A for an overview of the Pre-K Program). One significant difference in these two programs is that Pilot Pre-K sites may serve both 3- and 4-year olds, so students attending these programs may receive two years of Pre-K instruction. If feasible, analyses may be able to explore the relative impact of one year compared to two years of participation in Pre-K.

## Appendix A. Background and Implementation of Tennessee’s Pre-Kindergarten Program

Across the nation, access to high-quality state-funded Pre-Kindergarten (Pre-K) has steadily increased in the last 10 years. The National Institute for Early Education Research (NIEER) estimates that almost 1 million children participated in state Pre-K initiatives in 2005-2006, and spending in the states offering Pre-K totaled over \$3 billion.<sup>9</sup> Although state standards vary widely, more than three-quarters of state programs adhere to comprehensive early learning standards and more than half require teachers to have a Bachelor’s degree; 73% require teachers to have specialized Pre-K training. NIEER estimates 14% of 4-year-olds participated in state-funded Pre-K in 2002, but in 2006 20% of 4-year olds were enrolled.

The State of Tennessee has been funding Early Childhood Education (ECE) since the 1990s. Legislation enacted in 1996 permitted the creation of Pilot early childhood and Pre-Kindergarten programs for economically disadvantaged three- and four-year-olds. In the 1998-1999 school year, 30 Pilot Pre-K classrooms were created, serving approximately 600 students. Since then the program has grown to over 934 classrooms, serving approximately 17,000 children. Table A1 summarizes the number of students served and the number of classrooms in operation in Tennessee since 1998-1999, according to Tennessee Department of Education, Office of Early Learning.

**Table A1. Number of Students Enrolled in Tennessee Pre-K, 1998-1999 to 2007-2008**

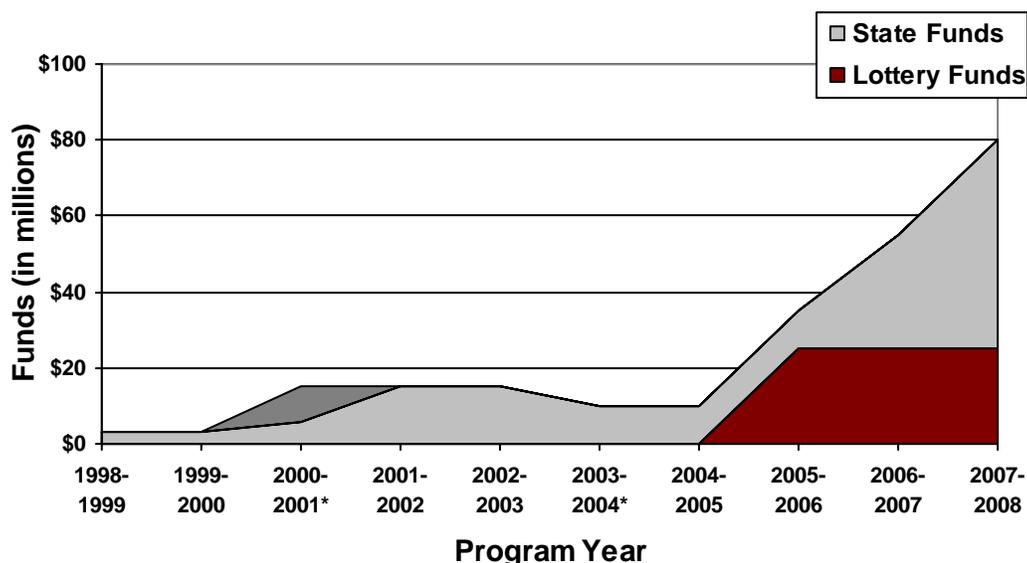
Program Year	Students Served	Number of Classrooms
1998-1999	600	30
1999-2000	600	30
2000-2001	3,000	150
2001-2002	3,000	90
2002-2003	3,000	150
2003-2004	2,900	150
2004-2005	2,900	147
2005-2006	8,900	446
2006-2007	13,000	677
2007-2008 (projected)	17,000	934

Source: State of Tennessee, Office of Early Learning, August 2007

The state Pre-K program has benefited from strong support from the Governor and bipartisan support in the Tennessee General Assembly. Together they passed the Voluntary Pre-K for Tennessee Act of 2005, increasing the state’s investment in Early Childhood Education and access for four-year-olds. The state allocated \$25 million from the excess net education lottery proceeds to fund approximately 300 new Pre-Kindergarten classrooms for at-risk four-year-olds—effectively tripling the number of students served. In the past three program years, state contributions have surged to \$55 million for a total of \$80 million for 2007-2008. Figure 1 shows the trends in funding over 10 program years.

<sup>9</sup> Barnett, W.S., Hustedt, J.T., Hawkinson, L.E., Robin, K.B. (2006). *The State of Preschool 2006*. National Institute for Early Education Research. Downloaded from the Internet at <http://nieer.org/yearbook/>.

**Figure A1. Tennessee Pre-K Funding by Source, 1999-2008**



Notes: \$9 million in funding was awarded and disseminated mid-year in program year 2000-2001 (dark gray area). TANF funding ended in 2003-2004 school year.

## Pre-K Pilot Sites

The expansion of the Pre-Kindergarten program in 2005 resulted in two systems of Pre-Kindergarten instruction: the Pilot programs that were begun in 1998 and the lottery/general fund-funded programs that were begun in 2005. The two systems are alike in their classroom requirements regarding teacher credentials, class size, and curricular focus; however, they differ in their funding amount and source and their affiliation with the public school system. Pilot Pre-K program sites were not required to be affiliated with a local education agency (LEA) although most were located in schools. Although the majority of Pilot Pre-K providers are LEAs, 14 private providers (approximately 22 classrooms) have continued to serve three- and four-year olds. In 2006-2007, approximately 3,000 students were served in the Pilot program.

In fiscal years 2002 and 2003, funding of the Pilot sites was supplemented by federal Temporary Assistance for Needy Families (TANF) funds. When the TANF funding ended in the 2003-2004 school year, the state became the program's sole funding source and each classroom's funding was reduced by \$30,000. In 2005-2006, each classroom received \$65,000 in state funds. A local match is not required; however, most pilot sites supplement the state allocation with local funds.

## Student Eligibility

Enrollment in the Voluntary Pre-K for Tennessee program is based upon legislation (TCA 49-6-101-104). The Pre-K state statute specifies that each LEA is authorized to and may enroll any at-risk child who is four years old by September 30 and resides in the geographic area served by the LEA. Priority is given to those children who are eligible for the

free/reduced price lunch (FRPL) program. The state of Tennessee Department of Education encourages school systems to accurately identify the number of unserved at-risk children in the school district, making every effort to fill Pre-K classrooms with at-risk children.

If, however, there is space available after priority is given to at-risk four-year-olds, the LEA may enroll students with disabilities, students identified as English language learners (ELLs), students in state custody, or students identified as educationally at-risk for failure due to circumstances of abuse or neglect.

At the end of the first Pre-K student attendance period, if an insufficient number of children meeting the aforementioned enrollment requirements are enrolled in a specific classroom, an LEA may submit a written request to the Office of Early Learning for approval to enroll children identified with other at-risk factors as determined by the local school board and the Pre-K Advisory Council, such as children with a parent(s) in the military deployed to active duty, teen parents, or parents with limited education. In these cases a written request must be submitted and approved by the Office of Early Learning.

If not enough at-risk children enroll to fill a classroom, students who do not meet any at-risk criteria but who are considered unserved or underserved may be enrolled after a written request is submitted and approved by the Office of Early Learning.

## **Classroom Requirements**

Classroom requirements for all Pre-K classrooms in Tennessee are the same for all sites, regardless of whether they are considered Pilot or Lottery/general fund-funded Pre-K programs. The requirements are as follows:<sup>10</sup>

- Maximum class size is 20 students.
- Each class must have at least one licensed teacher who is certified in early childhood education, and at least one educational assistant who has either a child development associate credential (CDA) or an associate degree in early childhood education, or who is working toward acquiring these credentials.
- The program must provide a minimum of five and one-half hours of quality instructional time per day.
- Classroom instruction must include the use of an educational, age-appropriate curriculum that aligns with the state department of education approved early learning standards and includes literacy, writing, math, and science skills.
- Instruction must also include a developmental learning program that addresses the cognitive, physical, emotional, social, and communication areas of child development.
- In addition, each program must comply with the state board of education's early childhood education and Pre-Kindergarten program rules and policies.

---

<sup>10</sup> Note: Complete scope of services for 2007-2008 (including program requirements) can be found on the State of Tennessee, Department of Education, Office of Early Learning web page:  
<http://state.tn.us/education/prek/documents/VoluntaryScopeofServices2007-08.doc>

## Program Effectiveness

Clearly, Tennessee's Pre-K program has experienced significant growth over the past three years. In 2005, the program served approximately 3,000 three- and four-year-olds in 148 classrooms funded with \$10 million in state revenue, but is expected to increase to over 900 classrooms (17,000 students) in 2007-2008, representing an increase of over 750 classrooms in just 3 years. By 2008, state Pre-K is expected to have expanded to 132 of the 136 school systems in 93 of the 95 counties in Tennessee. The Governor and Legislature have made Pre-K a priority in Tennessee, and funding for Pre-K education has increased from \$10 million in 2004-2005 to \$80 million in 2007-2008 through excess lottery funds and state revenue.

Collaboration is a distinctive characteristic of Tennessee's Pre-K program, as evidenced by the importance of classroom partnerships in the TN Pre-K program. In 2006-2007 there were 148 collaborative classroom partnerships between 39 local school systems and non-profit and for profit providers. Tennessee statute allows for state collaboration with agencies such as Head Start, Even Start, for-profit and not-for-profit child care providers, faith-based agencies, community-based agencies, and higher education institutions. Further, the program requires the active participation and collaboration of stakeholders at the local and state level in the form of Community and State Pre-K Advisory Committees.

The Tennessee Voluntary Pre-K program has been recognized as achieving 9 out of 10 quality standard benchmarks of the National Institute for Early Education Research (NIEER), for the past two years—one of only 6 states to achieve a score of 9 or 10.<sup>11</sup> These benchmarks include:

- Comprehensive early learning standards.
- Teacher degree of BA or higher.
- Teacher specialized training in Pre-K.
- Assistant teacher degree of CDA or equivalent.<sup>12</sup>
- Teacher in-service of at least 15 hours per year.
- Maximum class size of 20 students.
- Staff: child ratio 1:10 or better.
- Vision, hearing, health screenings and one support service.
- At least one meal per day.
- Monitoring/site visits.

As the program continues to grow and more children are exposed to high-quality early childhood education in Tennessee, research is increasingly able to investigate the short- and long-term impact of Pre-K on student outcomes in elementary and middle school. The state of Tennessee has been collecting data on student participation in Pre-K since the inception of the Pilot Pre-K program in 1998, and is in the unique position to track student outcomes longitudinally.

---

<sup>11</sup> Source: *The State of Preschool 2006*. National Institute for Early Education Research, [www.nieer.org](http://www.nieer.org).

<sup>12</sup> According to NIEER, Tennessee meets all these criteria with one exception (assistant teacher degrees). The full report is available for download from the Internet at <http://nieer.org/yearbook/>.

## Appendix B. Identifying Discrepancies in Gender and Race in Student Assessment Records

For any students with discrepant values for gender and/or race across sets of scores, if one value did not appear at least three times more often than any other value(s), all sets of scores for that student were flagged and excluded from analysis. If one value did appear at least three times more often than any other value(s), the scores for that value were retained as it was considered to be valid and only the set(s) of scores for the inconsistent value(s) was/were excluded.

Tables B1 and B2 below provide some hypothetical examples of students with discrepant values for gender and race. Please note that in both tables, “n/a” means the information is not available, either because the student does not have assessment scores for that grade/year or because a value was not entered for the student for that grade/year.

**Table B1. Hypothetical Examples of Students with Discrepant Values for Gender**

	Grade K 99-00	Grade 1 00-01	Grade 2 01-02	Grade 3 02-03	Grade 4 03-04	Grade 5 04-05
Student 1	Male	Male	Male	Female	Male	Male
Student 2	Female	Male	Male	n/a	n/a	n/a
Student 3	n/a	Female	Female	Female	Female	Male

Based on the hypothetical data in Table B1 above, the following actions would be taken:

Student 1: Only the student’s record for Grade 3 would be flagged and excluded from analysis.

Student 2: Records for all three grade/years would have to be excluded from analysis.

Student 3: Only the student’s record for Grade 5 would be flagged and excluded from analysis.

**Table B2. Hypothetical Examples of Students with Discrepant Values for Race**

	Grade K 99-00	Grade 1 00-01	Grade 2 01-02	Grade 3 02-03	Grade 4 03-04	Grade 5 04-05
Student 1	n/a	White	Black	White	White	White
Student 2	n/a	Asian	Black	Black	Black	n/a
Student 3	n/a	n/a	White	White	Asian	Asian

Based on the hypothetical data in Table B2 above, the following actions would be taken:

Student 1: Only the student’s record for Grade 2 would be flagged and excluded from analysis.

Student 2: Only the student’s record for Grade 1 would be flagged and excluded from analysis.

Student 3: Records for all four grade/years would have to be excluded from analysis.

## Appendix C. Imputing Missing Demographic Information

If a student's gender or race was not reported in a grade/year but was reported in another, the student's value for the missing grade/year(s) can be imputed from other records containing the missing information for the student. We used a strategy whereby a student's gender or race was imputed only if they had at least two other sets of assessment scores containing a value for gender or race—and, of course, it had to be the same value for all available sets of scores.<sup>13</sup>

Tables C1 provides hypothetical examples of student records with missing demographic information. Please note that in the table, "n/a" means a value was not entered for the student for that grade/year.

**Table C1. Hypothetical Examples of Students with Discrepant Values for Gender**

	Grade K 99-00	Grade 1 00-01	Grade 2 01-02	Grade 3 02-03	Grade 4 03-04	Grade 5 04-05
Student 1	Male	Male	Male	n/a	Male	Male
Student 2	Female	Male	Male	n/a	n/a	n/a
Student 3	n/a	n/a	Female	n/a	n/a	n/a

Based on the hypothetical data in Table C1 above, the following actions would be taken:

Student 1: The value "male" would be imputed into the student's records for Grade 3.

Student 2: No imputations can be done because of the discrepancies in the existing information.

Student 3: No imputations can be done because there is only one other set of scores for this student with a value for gender.

---

<sup>13</sup> Because Cohort 5 is comprised only of two grade/years (the Kindergarten class of 2003-2004 and the first grade class of 2004-2005), no imputations could be made for demographic characteristics.

## Appendix D. Pre-K Group Demographic Profile for Each Grade/Year

Because an increasing number of Pre-K students was assessed each year in Kindergarten, First, and Second Grade, the sample size and demographic characteristics also changed for each cohort and grade level. To ensure that the comparison group of non-Pre-K participants matched the Pre-K group as closely as possible for all analyses, separate stratified random samples were drawn for each cohort/grade level.

Table D1 provides a profile of the demographic characteristics for each grade level within each cohort.

**Table D1. Number and Percentage of Pre-K Students by Key Demographic Variables**

	FRPL				Non-FRPL			
	White		Non-White		White		Non-White	
	Male	Female	Male	Female	Male	Female	Male	Female
<b>Cohort 1</b>								
Grade K in 99-00	--	--	--	--	--	--	--	--
Grade 1 in 00-01	30.2% (16)	15.1% (8)	11.3% (6)	32.1% (17)	1.9% (1)	7.5% (4)	1.9% (1)	0.0% (0)
Grade 2 in 01-02	20.4% (23)	15.9% (18)	21.2% (24)	23.9% (27)	8.0% (9)	7.1% (8)	1.8% (2)	1.8% (2)
Grade 3 in 02-03	23.9% (34)	19.7% (28)	17.6% (25)	19.0% (27)	7.7% (11)	9.2% (13)	2.1% (3)	0.7% (1)
Grade 4 in 03-04	24.8% (35)	17.7% (25)	18.4% (26)	19.9% (28)	6.4% (9)	9.9% (14)	2.1% (3)	0.7% (1)
<b>Cohort 2</b>								
Grade K in 00-01	20.4% (10)	34.7% (17)	12.2% (6)	8.2% (4)	14.3% (7)	6.1% (3)	2.0% (1)	2.0% (1)
Grade 1 in 01-02	13.2% (20)	21.2% (32)	24.5% (37)	24.5% (37)	6.0% (9)	8.6% (13)	2.0% (3)	0.0% (0)
Grade 2 in 02-03	13.1% (31)	19.0% (45)	22.8% (54)	24.5% (58)	8.4% (20)	5.9% (14)	3.8% (9)	2.5% (6)
Grade 3 in 03-04	13.5% (34)	19.5% (49)	17.9% (45)	26.3% (66)	10.8% (27)	6.8% (17)	3.2% (8)	2.0% (5)
<b>Cohort 3</b>								
Grade K in 01-02	27.1% (16)	18.6% (11)	15.3% (9)	15.3% (9)	8.5% (5)	15.3% (9)	0.0% (0)	0.0% (0)
Grade 1 in 02-03	21.4% (74)	20.2% (70)	18.5% (64)	19.1% (66)	7.5% (26)	11.3% (39)	0.9% (3)	1.2% (4)
Grade 2 in 03-04	21.4% (90)	18.8% (79)	18.3% (77)	20.5% (86)	7.1% (30)	10.5% (44)	1.4% (6)	1.9% (8)
<b>Cohort 4</b>								
Grade K in 02-03	36.4% (64)	30.1% (53)	11.4% (20)	9.7% (17)	5.1% (9)	2.8% (5)	3.4% (6)	1.1% (2)
Grade 1 in 03-04	23.1% (173)	23.9% (179)	20.4% (153)	17.0% (127)	6.4% (48)	6.1% (46)	2.1% (16)	0.9% (2)

	FRPL				Non-FRPL			
	White		Non-White		White		Non-White	
	Male	Female	Male	Female	Male	Female	Male	Female
<b>Cohort 5</b>								
Grade K in 03-04	28.3% (72)	33.1% (84)	8.3% (21)	9.4% (24)	10.2% (26)	8.3% (21)	1.2% (3)	1.2% (3)

Great care was taken to ensure that the Pre-K sample matched the non-Pre-K samples in terms of gender, race, and school system (LEA) as closely as possible. Chi-square tests were conducted to compare the groups to make sure the proportions of subgroups were equivalent

The only significant differences between the Pre-K groups and the corresponding non-Pre-K samples were in the race distribution of three of the samples. In Cohort 2-Kindergarten and First Grade and in Cohort 3-Kindergarten the non-Pre-K sample contained slightly fewer cases in a race category combining Hispanic, Asian/Pacific Islander, and American Indian/Native American students than the Pre-K group. This is due to the generally small sample sizes associated with these minority groups. However, once the race measure was collapsed further to only two categories, white and non-white, there were no significant differences between the Pre-K group and the non-Pre-K sample.

The non-Pre-K samples were drawn from the three-category classification of race to better reflect the true distribution of the race of the Pre-K students in the non-Pre-K sample. However, there were too few cases in the Hispanic/Asian/Native American category to allow these cases to be properly analyzed. Thus they were combined with the Black students to create the minority group “non-white” for analysis.