
What Is the Penny Buying for South Carolina?



**Twentieth Annual Reporting on
the South Carolina Education
Improvement Act of 1984
Child Development Programs for
Four-Year-Olds:
Longitudinal Studies of Later
Academic Achievement**



South Carolina State Board of Education

December 2004

What Is the Penny Buying for South Carolina?

**Child Development Programs for Four-Year-Olds:
Longitudinal Studies of Later Academic Achievement,
1995–96 through 1999–2000 and
2000–01 through 2001–02**

**Prepared by the
Evaluation Section, Office of Research
South Carolina Department of Education**

**Inez M. Tenenbaum
State Superintendent of Education**

December 2004

Contents

General Introduction	1
Executive Summary	2
Section 1: A Longitudinal Study of Later Academic Achievement, 1995–96 through 1999–2000	
Review of the Literature	4
Study Design and Methodology.....	7
Population and Sample	7
Data Utilized	8
Data Analyses	8
Limitations to the Study.....	8
Research Questions.....	9
Findings.....	10
Conclusions.....	18
Recommendations.....	19
Section 2: A Longitudinal Study of Later Academic Achievement, 2000–01 through 2001–02	
Review of the Literature	21
The Purpose of the Study.....	27
Study Design and Methodology.....	27
Population and Sample	27
Data Utilized	28
Data Analyses	29
Limitations to the Study.....	29
Research Questions.....	30
Findings.....	31
Implications and Conclusions.....	39
Recommendations.....	41
Appendix A: Districts Providing MAT7 Test Data, Spring 2002	42
Appendix B: Summary of Longitudinal Studies on South Carolina Early Childhood Development Programs for Four-Year-Olds, 1986–2004	43
References.....	46

General Introduction

The Education Improvement Act (EIA) was one of the first in a series of education reform initiatives enacted by the South Carolina General Assembly. Ratified in 1984, the EIA introduced programs to recognize and foster superior student performance, identify and improve poor performance, and enhance student achievement. The General Assembly supported this reform effort through an increase of one cent in the state sales tax.

One of the programs funded by the EIA was the half-day child development program for four-year-old children with predicted readiness deficiencies that place them at risk for early school failure. The Target 2000—School Reform for the Next Decade Act of 1989 and the Early Childhood Development and Academic Assistance Act of 1993 (Act 135) expanded the scope of child development programs to include parent education and family literacy initiatives. The federal Title I program and the state-sponsored First Steps program provided additional funding.

The EIA requires the State Board of Education to submit an assessment of the half-day child development program for four-year-olds to the General Assembly by the first of December every year. This year's assessment report focuses on the later academic achievements of students who participated in child development programs in South Carolina. This report is comprised of two distinct longitudinal studies: program participants were followed from preschool through third grade in the first study, and the same cohort group was tracked from fourth to fifth grades in the second study. Wei Yao, of the Office of Research at the State Department of Education (SDE), served as the principal investigator on each of the studies.

Section 1 of this report is extracted from the 2002 edition of *What Is the Penny Buying for South Carolina?* This section presents the results of a longitudinal study comparing the academic achievement of child development program participants with that of their nonparticipant peers. Children who were in the four-year-old child development program during the 1995–96 school year were tracked to the third grade in the 1999–2000 school year.

Section 2 presents the continued tracking of the same cohort group of 1995–96 child development program participants through fourth and fifth grades in the 2000–01 and 2001–02 school years, comparing their academic achievement with that of their nonparticipant peers at grades four and five.

Executive Summary

Half-day child development programs, also known as four-year-old kindergarten or 4K, serve children from disadvantaged families—children who typically are not well prepared for school and are at risk of falling further behind as they continue their schooling. Without an appropriate education, these children have difficulty breaking the cycle of poverty when they reach adulthood. Priority for admission into these state-funded programs goes to children with backgrounds that historically have predicted significant readiness deficiencies. Criteria for enrollment include risk factors such as poverty (qualification for free- and reduced-price lunch), parents who have not graduated from high school, and various developmental delays. School districts use a diagnostic assessment to screen these children, and those with the lowest scores on the assessment are admitted to child development programs.

Since 1984, South Carolina has funded a half-day child development program to serve children who have been assessed as being at risk for academic failure. Legislation in 1986 and 1993 expanded the scope of child development programs to include parent education and family literacy initiatives. The federal Title I program and the state-sponsored First Steps program provided additional funding.

The Department of Education’s Office of Research followed children who attended four-year-old child development programs in 1995–96 and qualified for the free- or reduced-price school lunch program in the first through fifth grades. The achievement test scores of these children were compared to the test scores of a randomly selected group of children who did not participate in a child development program at age four but were similar to the participant group in that they also qualified for the free- or reduced-price school lunch program in the first grade. At the end of the fifth grade 3,486 child development program participants and 4,618 nonparticipants remained active in the study.

Findings indicated that, when compared to students similar in economic background, the overall benefits to four-year-old child development students appear to last through their early school grades. The following are the results at each grade level:

At the first grade:

- Child development program participants scored significantly higher than nonparticipants on school readiness as measured by the Cognitive Skills Assessment Battery (CSAB).
- When compared by gender, race, and eligibility for the free- or reduced-price lunch program, the CSAB school readiness mean scores of participants were significantly higher in all categories than those of nonparticipants.

At the second grade:

- Since no statewide assessment is conducted in the second grade, the Office of Research matched a small number of participants and nonparticipants to MAT7 (Metropolitan Achievement Tests, Seventh Edition) test scores provided by nine school districts. Although participants outperformed nonparticipants in all demographic categories, only the difference in the mean scores for mathematics was statistically significant.

At the third grade:

- Child development program participants scored significantly higher than nonparticipants on the PACT English language arts (ELA) and mathematics tests.

- When compared by gender and race, the mean scores for program participants were significantly higher than for nonparticipants in both ELA and mathematics. For students eligible for the free- or reduced-price lunch program, mean scores for program participants were higher for both ELA and mathematics, but the difference was statistically significant only for ELA.

At the fourth grade:

- Child development program participants scored significantly higher than nonparticipants on both PACT ELA and mathematics.
- When compared by gender, race, and eligibility for the free- or reduced-price lunch program, the PACT mean scores of program participants were significantly higher than those of nonparticipants in all categories.
- When comparisons were made on the proportion of students who scored at or exceeded the “proficient” level on the PACT, 23.9 percent of program participants and 23.2 percent of nonparticipants scored “proficient” or “advanced” on ELA. For mathematics, 15.6 percent of program participants and 15.2 percent of nonparticipants scored “proficient” or “advanced.”
- A significantly smaller proportion of program participants took PACT off-grade-level tests than did nonparticipants.

At the fifth grade:

- Child development program participants scored significantly higher than nonparticipants on both PACT ELA and mathematics.
- When compared by gender, race, and eligibility for the free- or reduced-price lunch program, the PACT mean scores of program participants were significantly higher than those of nonparticipants in all categories.
- When comparisons were made on the proportion of students who scored at or exceeded the “proficient” level on the PACT, 13.3 percent of program participants and 12.3 percent of nonparticipants scored “proficient” or “advanced” on ELA. For mathematics, the difference was statistically significant, with 17.7 percent of program participants compared to 15.6 percent of nonparticipants scoring “proficient” or “advanced.”
- A significantly smaller proportion of program participants took PACT off-grade-level tests than did nonparticipants.

Program participation also appeared helpful in significantly narrowing the performance gap between Caucasian and non-Caucasian child development program participants. The difference in the average PACT score between these groups for students who did not participate in a child development program was greater than that for the program participants. For program participants at the fourth grade, the difference in the average PACT score between the two racial groupings was 4.8 points for ELA and 4.9 for mathematics, compared to a gap of 8.0 points for ELA and 9.7 points for mathematics for the students who had not participated in child development programs. For the fifth grade the difference in average PACT scores for program participants was 6.5 points for ELA and 8.5 points in mathematics. For nonparticipants the gap was 10.8 points for ELA and 14.1 points for mathematics. For the program participants, the gap between the racial groups was statistically insignificant except for fifth-grade mathematics. The gap between the racial groups for nonparticipants was statistically significant for both ELA and mathematics at both the fourth and fifth grades.

SECTION 1

A Longitudinal Study of Later Academic Achievement, 1995–96 through 1999–2000

This longitudinal study of academic achievement attempts to evaluate the effects of participating in South Carolina's child development programs on children's later academic achievement. All child development participants and a randomly selected comparable group of nonparticipants were tracked from age four through the first three years of elementary school, with 15,143 students remaining active in the study (85 percent of the original subjects). Statistical analysis methods such as t-tests, analysis of variance (ANOVA), and analysis of covariance (ANCOVA) were utilized at the significance level of .05. Research questions addressed the differences in test performance between those students who participated in a child development program at the age of four and those students who did not. Test scores also are compared by demographic variables in an attempt to determine which groups of children benefited most from participation in a child development program.

The data analysis indicates that participating in the child development program for four-year-olds helps at-risk children perform significantly better academically at grades one and three compared to similar nonparticipants. Asian, Caucasian, and female students benefited more from the program than did others. Limited data from the nine districts (Aiken, Bamberg Two, Barnwell Forty-Five, Clarendon Two, Dillon One, Kershaw, Orangeburg Five, Saluda, and Williamsburg) revealed that although participants scored higher on the MAT7 (Metropolitan Achievement Tests, Seventh Edition) reading and mathematics at grade two, the difference was not enough to reach statistical significance. Program participation reduced students' need to spend more time per week being given academic assistance at grades two and three. Comparisons between full-day and half-day participants showed no significant differences on later academic performance.

Review of the Literature

Quality early childhood intervention programs are seen as one of the most effective ways to prevent learning difficulties and to promote healthy development and well being, especially among children from disadvantaged families (Reynolds 2002). Many researchers indicate that the negative effects of poverty can be reduced when children participate in high-quality early childhood educational programs (Schweinhart, Barnes, and Weikart 1993, Schweinhart 2001). Studies and discussions focus on issues such as whether programs had impact on the participants' later achievement and, if so, how long the program impact lasted. Who could benefit most from child development programs? What types of programs worked best for children who needed assistance to achieve success in school?

There appears to be little dispute about whether programs serving disadvantaged children have immediate or short-term effects (Barnett 2002). Repeatedly, studies have reported that, in the short-term, children in well-implemented intervention programs consistently show higher levels of cognitive development, early school achievement, and motivation than do children who do not participate in such programs. Barnett for example, in his widely cited study on the effects of

preschool programs (1995), synthesized ten studies that reported IQ gains at some point during or after program participation. In most instances, the gains were sustained through school entry at age five (Reynolds 2002).

However, there are different opinions on the persistence of long-term program effects. Barnett (1995 and 2002) examined a total of forty-three published research studies on large-scale public programs serving economically disadvantaged children at age four or younger. These studies measured participants' later achievement on at least one aspect of cognitive development, school progress, or socialization up to the third grade or later. Most of these studies utilized nontreatment comparison groups that were similar to the groups of children who participated in the intervention. No random assignment or quasi-experimental designs were used since such designs are often not practical in educational studies. The sample sizes in these studies ranged from 61 to 3,980 subjects.

Barnett (1995) found that long-term effects measured by achievement test results for reading and mathematics in large-scale programs were quite variable. Four of twenty-one studies of large-scale programs found no effects at any time. Five studies found initial effects that faded and ceased to be statistically significant by the end of the third grade. The other studies found statistically significant positive program effects in the third grade or later. The variation in findings with respect to the impact on long-term achievement could be the result of the quality of program implementation, design variations, high attrition of subjects, the lack of uniformity in the tests used to measure the achievements, or some other factors.

Other studies found positive results in both short- and long-term gains (Barnett 1995, Ramey and Ramey 1998). In his latest study, "Early Childhood Interventions: Knowledge, Practice, and Policy" (2002), Reynolds finds more evidence that program effects are significantly related to early and longer program participation, especially with regard to reading and mathematics achievement. Early participation provides greater learning opportunities for children when their cognitive, language, and motor skills are developing rapidly. Major studies of long-term programs, the Carolina Abecedarian Project (Campbell et al. 1998, Campbell and Ramey 1999) and the Chicago Child-Parent Centers (CPC) longitudinal research (Reynolds 2002), have also noted the advantage of early intervention.

The Carolina Abecedarian study reported a long-lasting benefit for children born to low-income families who were enrolled in an experimental early education program. Of the 111 children studied, 57 were continuously enrolled from infancy through age five in a high-quality early childhood program that used learning games to enhance children's abilities. The other 54 children who constituted the control group did not receive services. Researchers followed these children until age twenty-one. At that age, those students who had received early intervention were more likely to score higher on reading and mathematics tests, to be enrolled in or to have graduated from a four-year college, to have delayed parenthood, and to be gainfully employed (Campbell et al. 1998).

The CPC program is a center-based early intervention effort that provides comprehensive educational and family support services to economically disadvantaged children from preschool to the early elementary grades. The central goal of the program is, in the words of one authority,

“to reach the child and parent early, [to] develop language skills and self-confidence, and to demonstrate that these children, if given a chance, can meet successfully all the demands of today’s technological, urban society” (cited in Reynolds 2002, 114).

The longitudinal study of the CPC program included 989 low-income, mostly African American children who entered the program in preschool and finished kindergarten in 1986 and 550 children from similar disadvantaged neighborhoods who participated in an alternative all-day kindergarten program in the Chicago schools. The groups were well matched according to their eligibility for intervention, family socioeconomic status, gender, and race. At age twenty in 2000, 1,281 children (83 percent of the original sample) remained active in the study.

The CPC results presented clear evidence that participants were more ready to learn than were children who did not participate, and relatively high proportions of students in the experimental group scored at or above national norms on standardized tests. These effects carried over to later school achievement. For example, when they had reached the age range between eighteen to twenty years old, participating subjects were 29 percent more likely than those in the comparison group to have completed school. In addition, they had a 33 percent lower rate of juvenile arrest and a 40 percent lower rate of special education placement and grade retention (Reynolds 2002).

Researchers also note that children who have the same experiences in early education may vary in their later academic achievements (Barnett et al. 1987, Barnett 1995, Reynolds 2002). Children who were from disadvantaged families or who were from high-poverty neighborhoods benefited more from early interventions. There were some indications that boys benefited more from preschool, but girls benefited more from follow-up intervention (Reynolds 2002). Another recent national longitudinal study on the effects of early education demonstrated that children who lacked a positive learning environment did not achieve as well as those who did have those resources. These influences persisted from kindergarten through the first grade (Denton and West 2002).

Some programs appear to be more effective than others. Barnett (2002) discovered that school educational interventions (mostly part-day) for four-year-old disadvantaged children, including Head Start and public school programs, have larger estimated effects than child day-care programs. However, he warned that some caution should be exercised in drawing conclusions because programs vary with respect to the children served as well as in the research design adopted by the researchers.

What do we already know about the program effects of the child development programs for four-year-olds in South Carolina? In terms of short-term effects, a report on South Carolina preschool programs published in 1987 demonstrated that program participation helped program participants on performance measures at grade one (Barnett et al. 1987). This study followed a state sample of 362 preschool program (the 1983–84 cohort group of the child development programs for four-year olds) participants and 1,662 nonparticipants to grade one. It found that the students who participated in the preschool program were more likely to score above the readiness cut-off score on the CSAB (Cognitive Skills Assessment Battery) and were more likely to score higher on the first-grade BSAP (Basic Skills Assessment Program) reading and mathematics tests than were nonpreschool-program participants, though no statistically significant differences between groups were found.

More recently, three large-scale longitudinal studies conducted by the SDE's Office of Research provided additional evidence for the extended effects of early childhood programs. (The results of these studies were published by the SDE in 1998, 1999, and 2000 in successive editions of *A Longitudinal Research Report on the Early Childhood Development Program*.) All three studies constructed comparison groups of children who were matched on the basis of free- or reduced-price lunch program eligibility. Large sample sizes were utilized (8,235, 8,987, and 9,701 subjects respectively stayed active in each study). These studies tracked participating children from half-day child development programs at age four to the first grade. Comparisons of the CSAB school readiness scores of participants and nonparticipants were made. In spite of their higher risk for school failure, program participants performed equally as well as nonparticipants when entering the first grade at public schools. Among program participants it appeared that females and Asian and Caucasian students outperformed their peers. Participants' socioeconomic status and mothers' educational levels were positively related with the students' school readiness at grade one.

In South Carolina, program-effect studies have been limited to half-day child development program participants since the EIA and Act 135 required districts to establish at least one half-day program for children at risk. Program-effect differences between full-day and half-day child development programs were not known at the time that these three longitudinal studies were conducted. A review of the studies published over the past decade found no research concerning the long-term effect on the achievement of child development program participants beyond the first grade.

Study Design and Methodology

Since students could not be randomly assigned to a treatment group, a quasi-experimental design was utilized in which similar comparison groups were established. A longitudinal match on the program participants from the four-year-old child development program through the third grade with demographics and academic achievement test scores was completed. The comparison group was constructed by randomly selecting children who were not participants in child development programs but who were comparable in essential characteristics (eligibility for the free- or reduced-price lunches) to the children who were participants.

Population and Sample

All children (9,977 valid records) who participated in child development programs in 1995–96 were followed longitudinally through the third grade. Children participating in child development programs were deliberately identified and recruited through a screening process utilizing the Developmental Indicators in the Assessment of Learning-Revised (DIAL-R), along with supplementary information about the child's family such as education and income level.

A nonparticipant comparable group was randomly selected (7,889) from students eligible for free- or reduced-price school lunch. Any comparison group will consist of students who, on the whole, have lesser degrees of risk for school failure than the child development program participants. The 1999–2000 PACT (Palmetto Achievement Challenge Tests) scores of 15,143 third-grade children (85 percent of the original subjects) were examined.

Data Utilized

Five data sources were utilized in the investigation:

- the statewide survey on programs for four-year-olds conducted by the SDE's Office of Research during the 1995–96 school year;
- statewide student information files from the 1997–98 school year through the 1999–2000 school year (SDE precode file);
- statewide CSAB testing in grade one in the 1997–98 school year;
- MAT7 testing of available students in nine school districts in spring 1999 (when the cohort group was in the second grade); and
- statewide PACT testing in spring 2000 (when the cohort group was in the third grade).

Data Analyses

First-grade CSAB scores, second-grade MAT7 scores in reading and mathematics, and third-grade PACT English language arts (ELA) and mathematics scores of program participants and nonparticipants were compared. In order to determine whether the mean scores of the participants and groups of nonparticipants were significantly different, t-tests were utilized. Analyses of variances (ANOVA) were applied for comparisons of three or more groups when subpopulations were examined. When control of the extraneous variables was desired, analysis of covariance (ANCOVA) was used where possible to achieve statistical control of the extraneous variables in order to reduce error caused by initial differences on participants' later academic performances. The level of statistical significance was set at a probability value of .05 as the threshold; a probability below this threshold ($P < .05$) indicates that a difference of this magnitude could happen by chance less than 5 percent of the time.

Limitations to the Study

When one is designing educational program evaluation studies, certain limitations are inherently imposed. For the following reasons, this study should be interpreted with caution.

- A major limitation relates to the fact that due to ethical and practical considerations, individuals were not randomly assigned to treatment groups. With this limitation in mind, it is obvious that “true” experiments cannot be conducted when evaluating programs. This study employs quasi-experimental design; therefore, it is not feasible to completely rule out alternative explanations for the results.
- Uniform criteria for program implementation, instructional methods, the quality of teachers' professional development activities, and so forth were not mandated at the time when the data for this cohort group were collected.

- The comparison group for this study was randomly selected from nonprogram students eligible for free- or reduced-price school lunch. Students enrolled in the four-year-old child development programs typically have significant readiness-deficiency indicators other than low family income. Districts were required to identify and serve students at greatest risk for early school failure. Any selected comparison group likely will consist of students who, on the whole, have lesser degrees of risk.
- Statewide student achievement test scores were not available for grade two. The analysis in this report relied on a very limited sample of nine school districts that were not randomly selected, nor were the students guaranteed to be representative of all second-grade students in the state.
- The information about the nature of academic assistance provided to students in the primary grades using Act 135 funds was not available. Each school district could provide assistance uniquely to best serve local needs and considerations. Therefore, this study used only the number of hours that students received academic assistance.

For these reasons, statistical findings should be considered good but not exact. To maximize the internal and external validities, a comprehensive evaluation of the effectiveness and impact of child development programs for four-year-olds will require resolution of the above-described flaws in design and implementation.

Research Questions

1. What were the differences in student academic performance from grade one through grade three between child development program participants and nonparticipants? Did academic performance differ by demographic characteristics? (statewide first-grade CSAB data in 1997–98, nine districts’ second-grade MAT7 data in spring 1999, and statewide third-grade PACT data in spring 2000)
2. Among child development program participants, which group of children benefited most in terms of academic achievement up to the third grade? (data application same as question 1)
3. Did continued academic assistance given to program participants from grade two through grade three affect their academic performance? (second-grade MAT7 data from nine school districts in spring 1999 and statewide third-grade PACT data in spring 2000)
4. What were the academic performance differences between child development program participants in half-day programs and those in full-day programs in grades one through three? Did their performances differ by demographic features? (data application same as question 1)

Findings

1. What were the differences in student academic performance from grade one through grade three between child development program participants and nonparticipants? Did academic performance differ by demographic characteristics?

Students who participated in child development programs for four-year-olds scored higher on first-grade school readiness assessments, second-grade MAT7 tests, and third-grade PACT tests.

First Grade

The comparisons were conducted on first-grade school readiness scores between child development program participants and comparable nonparticipants randomly selected from students who were eligible for the free- or reduced-price school lunch program.

- Students who participated in the child development programs for four-year-olds scored significantly higher in school readiness as measured by the CSAB at grade one.
- In comparisons between the student demographic features of participants and nonparticipants, the findings significantly favored the program participants among Caucasian, non-Caucasian, male, female, and students eligible for the free- or reduced-price lunch program at school. Table 1 describes the details.

TABLE 1

**Comparison between Child Development Program Participants and Nonparticipants:
Mean CSAB School Readiness Scores by Demographic Category, Fall 1997**

Demographic Category	Program Participants N=7,515	Nonparticipants N=7,889	Probability Value
All	93.9	92.1	< .05
Male	92.9	90.6	< .05
Female	94.7	93.2	< .05
Caucasian	95.6	93.8	< .05
Non-Caucasian	92.5	91.3	< .05
Eligible for free- or reduced-price lunch	92.4	92.1	< .05

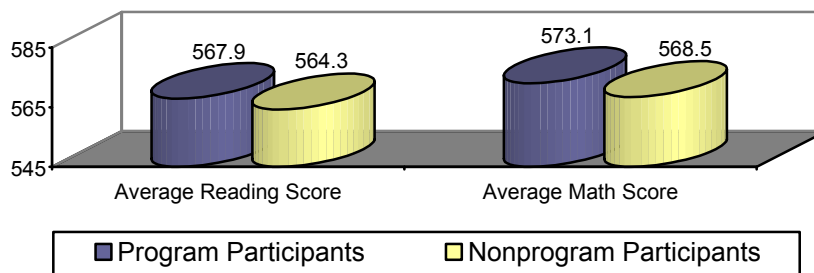
Second Grade

The analysis utilized MAT7 test data available from nine school districts. Child development program participants were matched to the second-grade MAT7 data (N=1,224). The same randomly selected nonprogram students eligible for the free- or reduced-price lunch program who served as the first-grade comparison group were matched to their second-grade (N=711) MAT7 test scores in the nine school districts for performance comparisons.

- Child development program participants in the nine school districts scored higher on the second-grade MAT7 tests in reading and mathematics than nonparticipants. However, the gaps were not large enough to be statistically significant. Figure 1 shows the differences.

Figure 1

Comparison between Child Development Program Participants and Nonparticipants: Second-Grade MAT7 Performance in Nine School Districts, Spring 1999



- In comparing test scores between participants and nonparticipants, the findings favored the program participants among Caucasian, non-Caucasian, female, and students eligible for free- or reduced-price lunch—though the difference was not large enough to be statistically significant. Only male participants scored significantly higher than the male nonparticipants on mathematics. Table 2 gives the details.

TABLE 2

Comparison between Child Development Program Participants and Nonparticipants: MAT7 Reading and Mathematics Performance by Demographic Category, Spring 1999

Demographic Category	Mean Scores Reading		Mean Scores Math		Probability Value	
	Program N=1,213	Nonprogram N=696	Program N=1,216	Nonprogram N=708	Reading	Math
Male	566.7	561.0	571.6	565.2		< .05
Female	570.4	569.5	575.8	574.1		
Caucasian	569.6	564.4	574.7	570.2		
Non-Caucasian	564.0	562.4	569.4	566.0		
Eligible for free- or reduced-price lunch	564.1	563.1	569.9	567.9		

Third Grade

About fifteen thousand students (15,143), 85 percent of the original subjects statewide, were matched from prekindergarten at age four through the third grade on the basis of PACT scores in 2000. The PACT scores of participants and those of randomly selected nonparticipants comparable in eligibility for the free- or reduced-price lunch program were compared.

- Child development program participants scored significantly higher on the third-grade PACT in spring 2000 on both subject areas of ELA and mathematics than nonparticipants ($P < .05$). Figure 2 shows the performance comparisons.
- With the exception of students from economically disadvantaged families, all subgroups of program participants (based on the demographic characteristics—male, female, Caucasian, non-Caucasian) scored significantly higher in both subject areas on the PACT than nonparticipants ($P < .05$). The higher scores on mathematics favored child development program participants eligible for the free- or reduced-price lunch program, but the difference was not large enough to be statistically significant (see table 3).

Figure 2
Comparison between Child Development Program Participants and Nonparticipants:
Third-Grade Performance on the PACT, Statewide, Spring 2000

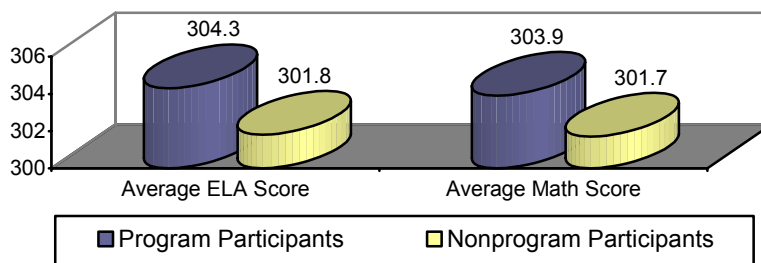


TABLE 3

Comparison between Child Development Program Participants and Nonparticipants:
Third-Grade Performance on the PACT by Demographic Category, Spring 2000

Demographic Category	Mean Scores ELA		Mean Scores Math		Probability Value	
	Program N=7,298	Nonprogram N=7,658	Program N=7,342	Nonprogram N=7,725	ELA	Math
Male	302.9	299.8	304.2	301.2	< .05	< .05
Female	305.6	303.9	303.6	302.2	< .05	< .05
Caucasian	309.0	307.9	309.4	308.3	< .05	< .05
Non-Caucasian	299.8	298.2	298.9	297.8	< .05	< .05
Eligible for free- or reduced-price lunch	300.1	298.9	299.1	298.8	< .05	

2. Among child development program participants, which group of children benefited most in terms of academic achievement up to the third grade?

Asian and Caucasian participants consistently outperformed African American program participants from grade one to grade three. Participating children not eligible for the free- or reduced-price school lunch program demonstrated higher gains.

The program's impact on groups of participants was measured by comparisons of their scores on the first-grade CSAB, the second-grade MAT7 reading and mathematics tests, and the third-grade PACT ELA and mathematics assessments. Comparisons were made between subgroups by gender, race, and family income (as measured by eligibility for the free- or reduced-price lunch program). Any demographic groups of participating children with fewer than thirty students were excluded. Tables 4 through 6 give detailed comparisons by gender, race, and lunch status, respectively.

TABLE 4
Later Academic Performance of Child Development Program Participants,
1997–98 through 1999–2000 School Years, by Gender

	Grade 1 School Readiness Mean Scores	Grade 2 MAT7 Mean Scores		Grade 3 PACT Mean Scores	
	Statewide	Nine School Districts		Statewide	
Gender	N=7,515	Reading N=1,194	Math N=1,197	ELA N=7,294	Math N=7,337
Female (F)	95.1	570.7	576.1	305.6	303.1
Male (M)	93.2	569.9	574.6	302.9	303.7
Gender group comparison results	F > M			F > M	
Probability value	< .05			< .05	

TABLE 5

**Later Academic Performance of Child Development Program Participants,
1997–98 through 1999–2000 School Years, by Race**

	Grade 1	Grade 2		Grade 3	
	School Readiness Mean Scores	MAT7 Mean Scores		PACT Mean Scores	
	Statewide	Nine School Districts		Statewide	
Race	N=7,515	Reading N=1,194	Math N=1,197	ELA N=7,294	Math N=7,337
Asian	97.3	573.1	577.9		
Caucasian	95.9			309.0	309.4
Other	94.2				
African American (AA)	92.8	567.3	572.6	299.8	298.8
Hispanic	91.2			309.0	307.3
Racial group comparison results	Asian > AA & Hispanic Caucasian > AA & Hispanic	Caucasian > AA	Caucasian > AA	Caucasian & Hispanic > AA	Caucasian & Hispanic > AA
Probability value	< .05	< .05	< .05	< .05	< .05

TABLE 6

**Later Academic Performance of Child Development Program Participants,
1997–98 through 1999–2000 School Years, by Lunch Status**

	Grade 1	Grade 2		Grade 3	
	School Readiness Mean Scores	MAT7 Mean Scores		PACT Mean Scores	
	Statewide	Nine School Districts		Statewide	
Lunch status	N=7,515	Reading N=1,194	Math N=1,197	ELA N=7,294	Math N=7,337
Not eligible for free- or reduced-price (NF/R) lunch	96.9	574.4	578.6	310.4	310.4
Eligible for free- or reduced-price lunch (F/R)	92.7	568.5	573.9	300.0	299.5
Lunch group comparison results	NF/R lunch > F/R lunch	NF/R lunch > F/R lunch		NF/R lunch > F/R lunch	NF/R lunch > F/R lunch
Probability value	< .05	< .05		< .05	< .05

Statewide data analyses comparing subgroups among program participants provided evidence that at grades one and three, Asian, Caucasian, and Hispanic children (third grade) and children not eligible for free- or reduced-price lunch scored higher on school readiness measures and on third-grade reading and mathematics. Female participants scored significantly higher on the CSAB and the third-grade PACT ELA. The second-grade data available from the nine school districts revealed that Caucasian participants had better performances than African Americans on the MAT7 reading and mathematics tests. Also the economic status of participants' families (as determined by lunch status) was positively related with second-grade MAT7 reading performances. Children not eligible for the free- or reduced-price lunch program outperformed those who were eligible for the program, except in mathematics at grade two.

More intensive data analyses using ANCOVA were conducted in order to reduce the error caused by initial differences among students when they were entering the program and at the first grade. Adjusted mean scores were obtained by removing initial differences at program entry on DIAL-R scores and differences in CSAB scores possibly caused by students' having participated in other preschool programs. Similar results were obtained from statewide data analysis except for participants' third-grade mathematics scores on the PACT. The adjusted mean scores of male participants were significantly higher than those of their female counterparts, while without statistical control on the extraneous variable, no gender differences were found. The second-grade performance analyses using participants from the nine districts revealed that after the initial differences were removed, participants from disadvantaged families (eligible for free- or reduced-price lunch program) performed equally well on the MAT7 reading test as students from families not identified as disadvantaged.

<p>3. Did continued academic assistance given to program participants from grade two through grade three affect their academic performance?</p>

Average grade-two MAT7 scores from nine school districts showed no difference between program participants receiving one to two hours of academic assistance and those participants receiving over two hours of academic assistance. Statewide, child development program participants who received academic assistance for one to two hours per week scored significantly higher on ELA and mathematics as measured by the third-grade PACT than participants who received assistance more than two hours per week.

Based on a sample of nine districts, at the second grade 27.2 percent of child development participants and 40 percent of nonparticipants received over two hours of academic assistance per week. At grade three, statewide the same percentage (31 percent) of child development program participants and nonparticipants received over two hours of academic assistance per week. These results indicate that the children who participated in the child development program, although judged at the age of four to be at risk for academic failure, were no more likely to need extra academic assistance at the third grade than their nonparticipant peers.

Child development program participants receiving academic assistance were divided into two groups in the analysis explained below, the group receiving one to two hours per week of academic assistance and the group receiving more than two hours per week. The comparisons measured the difference in mean test scores for child development program participants by the hours of weekly academic assistance received at grades two and three.

Second Grade

The analyses of data from the small sample of nine districts revealed that child development program participants' having received a greater amount of academic assistance (over two hours) weekly was not significantly associated with better MAT7 performances in reading or mathematics at grade two.

Third-Grade ELA and Mathematics

Data analysis found that, statewide, child development program participants receiving fewer hours of academic assistance (one to two hours per week) had significantly higher average scores in both ELA and mathematics on the PACT at grade three. Table 7 depicts the differences.

TABLE 7
Later Performances of Child Development Program Participants
by Hours of Academic Assistance Received Weekly
Grade 3, Statewide PACT Data, Spring 2000

	PACT ELA Mean Score	PACT Math Mean Score
Receiving 1–2 hours of academic assistance per week (1–2 hrs.)	304.0	303.5
Receiving more than 2 hours of academic assistance per week (over 2 hrs.)	302.3	302.6
Group comparison results	1–2 hrs. > over 2 hrs.	1–2 hrs. > over 2 hrs.
Probability value	< .05	< .05

Although the difference in average PACT scores by the hours of academic assistance that students had received was significant, the average PACT scores for child development program participants were higher than those of their nonparticipant peers, regardless of the number of hours of academic assistance (see figure 2, above).

4. What were the academic performance differences between child development program participants in half-day programs and those in full-day programs in grades one through three? Did their performances differ by demographic features?

First Grade

Statewide data showed that children from full-day and half-day programs scored very similarly on school readiness measured by the CSAB at grade one.

Due to the limited number of children (N=877) served by full-day child development programs, a random selection procedure was used to select 991 participants from half-day programs in order to make two similar-sized groups. The first-grade CSAB scores of half-day program

participants were compared to the scores of full-day program participants. The findings were that children from these two types of programs scored equally well on first-grade school readiness assessment. No statistical significance was found. Disaggregated student data were also examined between male and female, Caucasian and non-Caucasian, and students eligible for free- or reduced-price lunch and students paying full price for school lunch. No statistical differences were discovered in school readiness scores on the CSAB at grade one. Table 8 presents the details.

TABLE 8
Comparison between Full-Day and Half-Day Child Development
Program Participants at Grade One: CSAB School Readiness
Mean Scores by Demographic Category, Fall 1997

Demographic Category		Half-Day Program Participants N=991	Full-Day Program Participants N=877
All		93.9	93.8
Gender	Male	93.2	92.4
	Female	94.1	94.7
Race	Caucasian	95.0	96.0
	Non-Caucasian	92.6	92.4
Lunch status	Eligible for free- or reduced-price lunch	92.7	92.5
	Not eligible for free- or reduced-price lunch	95.6	96.6

Second Grade

Comparison was not meaningful as MAT7 test data available from nine school districts provided an insufficient number of full-day participants (N < 30) to yield statistically reliable information.

Third Grade

The program participants selected for study at grade one were tested at grade three. Participants from full-day programs and those from half-day programs scored equally well on third-grade ELA and mathematics.

There was no performance difference on the third-grade PACT performance between the full-day and half-day child development program participants. When disaggregated data were analyzed by male, female, Caucasian, non-Caucasian, eligibility for the free- or reduced-price lunch program, and fully paid lunch students, the differences found were too small for statistical significance. Table 9 shows the comparisons of third-grade PACT scores between full-day and half-day program participants.

TABLE 9

**Comparison between Full-Day and Half-Day Child Development Program Participants:
Third-Grade PACT Performance by Demographic Category, Spring 2000**

Demographic Category		ELA		Math	
		Mean Scores	Mean Scores	Mean Scores	Mean Scores
		Full-Day N=857	Half-Day N=951	Full-Day N=867	Half-Day N=957
All		304.9	304.4	303.9	303.1
Gender	Male	302.4	302.7	304.1	303.6
	Female	302.4	302.7	304.1	303.8
Race	Caucasian	308.9	308.3	310.0	308.4
	Non-Caucasian	299.7	298.1	299.0	297.3
Lunch Status	Eligible for free- or reduced-price lunch	299.0	297.5	298.8	297.0
	Not eligible for free- or reduced-price lunch	310.3	309.9	311.1	309.7

Conclusions

This study focused on the later academic performances of children who participated in the 1995–96 class of the child development program for four-year-olds. It followed the cohort group for four years and compared student performances on the CSAB first-grade school readiness assessment, the second-grade MAT7 tests, and the third-grade PACT assessment between participants and nonparticipants as well as within-program disaggregated populations. It also investigated the issue of whether the different hours of academic assistance that participants received per week from grade two to grade three helped their performance. Finally, it looked into the performance differences between full-day and half-day child development program participants on the same three tests (named above) from grade one through grade three. Detailed data analyses yielded the following conclusions:

- Child development programs for four-year-olds had a positive long-term effect on participants’ later academic performances in comparison to similar students who did not participate in the program. By definition, the majority of program participants were children whose developmental indicators—including their families’ economic and educational backgrounds—placed them at risk academically. In spite of their risk levels, the program participants statewide demonstrated significantly higher scores than nonparticipants on the CSAB first-grade readiness assessment and the third-grade PACT in ELA and mathematics. Comparisons on the second-grade MAT7 reading and mathematics tests revealed that participants scored higher than nonparticipants, but the difference was not large enough to be statistically significant. When disaggregated data analyses were conducted by male, female, Caucasian, non-Caucasian, and free- or reduced-price school lunch students, participants in all subgroups scored significantly higher on the CSAB than nonparticipants. At grade two,

only male participants scored significantly higher than their nonparticipant peers on mathematics. At the third grade, all participants belonging to the above-mentioned demographic groups scored higher in both ELA and mathematics on the PACT than nonparticipants—except for those participants eligible for the free- or reduced-price lunch program, who scored higher than their nonparticipant peers only on the ELA section of the PACT.

- It appears that child development programs helped Asian, Caucasian, and Hispanic children more than other ethnic groups of children in their later achievement performances on reading and mathematics. Female participants benefited more than male participants from the child development program on measures of reading. Child development program participation appeared less effective for African American students and participants eligible for free- or reduced-price lunch than for students in other demographic categories.
- Fewer child development program participants than nonparticipants needed academic assistance more than two hours per week at grade two. Participants who received one to two hours of academic assistance per week demonstrated significantly better performance on third-grade PACT ELA and mathematics than their counterparts who needed more than two hours of academic assistance per week.
- Statewide data analysis on the long-term program effect on academic performance of full-day and half-day program participants illustrated that full-day and half-day participants performed equally well on the CSAB and the third-grade PACT tests. Second-grade MAT7 student performance data for full-day and half-day participants were insufficient (from nine school districts) to provide statistically reliable results in a comparison of the two.

Recommendations

The following recommendations are based upon the study results:

- Studies on other program-effect indicators such as the reduction of the grade retention rate or the rate of placement of children in special education should be conducted to illustrate, in more depth, the positive effects that child development programs have on academic performances.
- Studies should be conducted to identify curricula and teaching methodologies that are more sensitive to the needs of African American participants and those eligible for the free- or reduced-price lunch program.
- The nature and quality of academic assistance provided to children with academic difficulties in primary grades need to be investigated to determine academic assistance methods more effective than merely additional time.
- Program impact on children’s cognitive and social development should be evaluated on the basis of both quantitative and qualitative data.

- An examination of the similarities and differences between full-day and half-day programs should be conducted to better understand the relationship between types of programs and participants' later performances.
- A study of the relationship between leadership expertise (that of directors or supervisors of the child development programs) and student outcome should be developed to further address the issue of program quality.
- A study of the relationship between participating children's later academic achievement and the curriculum or pedagogy used in classrooms should be carried out to provide empirical data for identifying effective instructional practices.
- Research on the effect of the duration of program participation on children's achievement at primary grades should be conducted. It will provide evidence to parents and educators regarding the minimum amount of time that children with predicted school readiness deficiencies should participate in child development programs.

Section 2

A Longitudinal Study of Later Academic Achievement, 2000–01 through 2001–02

This study examined the long-term program impact on later academic performance of children who participated in South Carolina’s child development programs at age four in school year 1995–96. It was a continuation of the longitudinal study conducted in 2002.

Program participants eligible for free- or reduced-price school lunch and a randomly selected comparable group of nonparticipants who also were eligible for free- or reduced-price lunch were tracked across six years through grade five. Comparisons focused on student performance between participants and nonparticipants at grades four and five as measured by the PACT in spring 2001 and 2002. Test performance disaggregated by student demographic characteristics and daily program length (full-day or half-day) were also examined. Further investigations were focused on the performance gap between Caucasian and non-Caucasian students who had participated in a child development program in comparison with the gap between Caucasian and non-Caucasian students who had not participated in such a program. As a final point, the study examined whether program participation helped reduce the number of students who had to take PACT off-grade-level tests. Statistical analyses of t-tests, ANOVA, and chi-square tests were used at the significance level of .05.

Key findings regarding students in grades four and five are as follows:

- Program participants scored significantly higher than nonparticipants in both ELA and mathematics as measured by the PACT.
- Participation in child development programs significantly helped reduce the performance gap between Caucasian and non-Caucasian students.
- The proportion of program participants taking PACT off-grade-level tests was significantly smaller than nonparticipants.

No significant difference in effect was found between full-day and half-day programs.

Review of the Literature

Educators know that children from disadvantaged families typically are not prepared for school and that these children risk falling further behind as they continue through school. Without an appropriate education, these children may never be able to break the cycle of poverty when they become adults. Research shows that the most effective intervention for children at risk of school failure begins in preschool (Barnett 1995; Schweinhart, Barnes, and Weikart 1993; Reynolds 2001). Researchers, policy makers, and educators have now turned their attention to the duration of preschool programs’ impact in closing the achievement gap between children living in poverty and their more affluent peers.

Many longitudinal studies have been conducted on the long-term effects of preschool programs enrolling children who are at risk for school failure. Findings vary due to differences in program

structure, accessibility, duration, classroom characteristics, comprehensiveness of services, and parent involvement. Studies on the impact of program intervention have focused on two types of preschool programs: the small-scale pilot and model programs and the large-scale, state-funded programs.

Studies of small-scale pilot and model preschool programs serving children at risk:

Researchers have conducted studies on the long-term program impact of preschool programs on participants' later success and provided solid evidences for the effectiveness of small-scale pilot and model programs. Two seminal programs they have examined are the High/Scope Perry Preschool Program and the Carolina Abecedarian Project.

The High/Scope Perry Preschool Study was begun in the early 1960s by David P. Weikart, a special education director in the public school system in Ypsilanti, Michigan. Dissatisfied with their school district's use of grade retention as a way of coping with rampant school failure, Weikart and his colleagues developed a program of preschool education—now known as the High/Scope model—and initiated a study to evaluate the effects of that program. Lawrence J. Schweinhart joined the evaluation team in 1975 (Schweinhart 2002). The High/Scope Perry Preschool Study, a pioneering systematic evaluation of the effect of preschool education, is one of the few longitudinal studies that examined programs involving the random assignment of children to either the treatment group or the control group.

Seeking to promote the social and cognitive development of at-risk children, the High/Scope Perry Preschool Program identified 123 African Americans born into poverty and at high risk of failing in school. Subjects were randomly assigned either to an initially similar treatment group who received a high-quality preschool education within the Ypsilanti public schools or to a nonprogram control group who received no preschool education. Data were collected on both groups from age three through age twenty-seven. The program participants significantly outscored the nonprogram participants on a test of general literacy at age nineteen as well as on sections of this test at age twenty-seven and on school achievement tests in reading, language, and arithmetic given at age fourteen. Positive long-term effects also were found in the program participants' economic success in early adulthood and in their having fewer criminal arrests than their nonprogram peers (Schweinhart, Barnes, and Weikart 1993). A cost-benefit analysis showed that the average annual cost of the program was \$14,716 per participant, with a cost-benefit ratio of 7.16 to 1 (Schweinhart 2003, 5).

Another exemplary program was the Abecedarian early childhood intervention project in North Carolina. The Abecedarian project was designed to study the impact of educational intervention in the lives of preschool-aged children from low-income families. In this program, children from birth to age five received an educational intervention; researchers measured cognitive achievement and other factors through a battery of standardized measures. Upon completion of the intervention, the children entered school; they were studied and tested again at ages twelve and fifteen years ("Carolina Abecedarian Project" 2003).

Designed as a randomized, controlled trial, the project was initiated in the summer of 1972 with a pool of 111 randomly selected infants and their families; 57 of the children were randomly assigned into the Abecedarian program, and the remaining 54 children were assigned to the control group. Both groups received nutritional supplements during the first years of life, and

social service referrals when needed throughout the first eight years of life. The 57 children received a carefully monitored educational intervention for the first five years of life—a year-round, all-day educational childcare/preschool program emphasizing the development of cognitive, language, and adaptive behavior skills. The 54 assigned to the control group received nutritional supplements in infancy and supportive social services. The control group received no educational intervention, however (“Carolina Abecedarian Project” 2003).

The treated and untreated children were initially comparable with respect to scores on infant mental and motor tests. However, from the age of eighteen months and through the completion of the child-care program, the children in the intervention group had significantly higher scores on mental tests than those in the control group. Follow-up cognitive assessments completed at ages twelve and fifteen years showed that the intervention group continued to have higher average scores on mental tests. Effect sizes remained moderate. Treated children scored significantly higher on tests of reading and math from the primary grades through middle adolescence. Effect sizes for reading were large; those for math were large to moderate (Campbell and Ramey 1999). In comparison with their peers in the control group at age fifteen, the program participants had a lower rate of grade retention in kindergarten through grade nine, were less likely to need special education in kindergarten through grade nine, and had higher adjusted mean reading and mathematics scores on the Woodcock-Johnson test (Campbell et al. 2001).

Researchers followed these children to age twenty-one, at which point cognitive functioning, academic skills, educational attainment, employment, parenthood, and social adjustment were measured. Of the original 111 infants, 104 (53 from the intervention group and 51 controls) were assessed. Those who received early educational intervention had significantly higher mental test scores from toddlerhood through age twenty-one than did untreated controls. Averaged over the age span tested, the mental test score effect size for the treatment group was moderate but considered educationally meaningful. Enhanced language skills in the children appears to have mediated the effects of early intervention on mental test performance (i.e., cognitive skills). Reading achievement scores were consistently higher for individuals with early intervention (Campbell and Ramey 1999).

Young adults from the intervention group were significantly more likely still to be in school at age twenty-one—40 percent of the intervention group compared with 20 percent of the control group. A significant difference was also found for the percent of young adults who ever attended a four-year college. About 35 percent of the young adults in the intervention group had either graduated from or were at the time of the assessment attending a four-year college or university. In contrast, only about 14 percent in the control group had done so. Employment rates were higher (65 percent) for the intervention group than for the control group (50 percent), although the trend was not statistically significant (Campbell and Ramey 1999).

After conducting a benefit-cost analysis of the Abecedarian project, researchers from the National Institute for Early Education Research (NIEER) concluded that the average annual cost of the program was about \$13,000 per child in 2002 dollars. Although this is double the cost of the average Head Start program, the NIEER researchers found “benefits outweighed the costs by a factor of four dollars for every dollar spent” (NIEER 2004).

The results indicated that a high-quality program effect could last for more than a decade after the participants had left the program. The random assignment procedures used at the beginning

of these program interventions allowed investigators to conclude that the positive effects were the result of program participation because these experiments permitted a high degree of control of nuisance variables. However, as Schweinhart observes, “these studies prompt continued caution about generalizing their claims to all preschool programs. The High/Scope and Abecedarian programs were model programs, run under the watchful eyes and with the active support of scientists and expert program developers. . . . These were . . . adequately supported, professionally run programs; they do not generalize to programs that are not adequately supported and professionally run” (Schweinhart 2003, 7).

Studies of large-scale, state-funded preschool programs serving children at risk:

Nationwide, the number of state-funded preschool programs for at-risk children increased dramatically during the 1990s. By the end of 2003, thirty-nine states and the District of Columbia had designed, implemented, and funded their own prekindergarten programs on a large scale. Most state-funded preschool programs give enrollment priority to children whose families have low incomes or who are otherwise at risk for poor school achievement (U.S. Department of Health and Human Services 2003).

The large-scale, state-funded programs are different from model programs. For example, the random assignment of subjects into treatment or control groups is impractical in these large-scale programs. In addition, the states’ commitment to funding, implementing, and conducting a comprehensive preschool program may vary (Gilliam and Ripple 2004), and program implementation and support may differ from site to site. Moreover, programs and agencies that rely on public funds increasingly have been held accountable for demonstrating their effectiveness. Evaluation of program implementation and impact is mandated in the state legislation that authorizes the particular program. Investigations into any tangible long-term effects of large-scale, state-funded preschool programs have unfailingly attracted the attention of researchers, educators, and government policy makers.

Walter S. Gilliam and Edward F. Zigler, of the Yale University Child Study Center, conducted a thorough meta-analysis of all evaluations of state-funded preschool programs from 1977 to 1998. Of the thirty-two states funding preschool programs, only twelve (Arkansas, Florida, Georgia, Kentucky, Louisiana, Maryland, Michigan, North Carolina, South Carolina, Texas, Vermont, and Washington) had completed or were conducting impact evaluations of their programs. The District of Columbia had also funded and evaluated a preschool program. Third-party evaluators had conducted seven of these thirteen program-impact evaluations, and state departments of education had conducted six. All but three of the evaluations used some form of comparison group against which program impacts were estimated. No evaluation randomly assigned children to program and control groups. The median length of follow-up was to the third grade. The number of subjects used in the evaluations varied significantly by state, based on the evaluation method used. The sample size ranged from fourteen matched pairs to over 40,000 children per group. Most study attrition rates ranged from less than 10 percent to about 25 percent per year (Gilliam and Zigler 2000, 5–8).

The twelve state and District of Columbia evaluation studies followed students in their later success in eleven domains, though the number of domains studied differed from state to state. The most commonly measured outcome domains were special education placement and academic achievement. A large number (forty-two) of different tests and procedures were used to

assess student academic achievement. Several of the tests were well-known, psychometrically valid instruments (the California Achievement Test, the Stanford Achievement Test, the Metropolitan Achievement Test). However, in many cases, relatively unknown tests—with little data provided regarding their reliability and validity—were used. Among the ten evaluations that used comparison groups to address the impact on academic performance, six states (Florida, Georgia, Maryland, New York, South Carolina, and Texas) and the District of Columbia reported statistically significant impacts on academic achievement test scores occurring at one or more grade levels (Gilliam and Zigler 2000, 14).

Maryland conducted the longest impact longitudinal evaluation known for a state-funded preschool program serving at-risk children, following 356 prekindergarten graduates and 305 nonprekindergarten students to the eighth grade. Results showed that prekindergarten participants were less likely than nonparticipants to be classified as at risk, to be assigned to special education classes, or to be retained. They also scored significantly better than nonparticipants on the reading, language arts, and mathematics portions of the CAT (California Achievement Test) in grades three, five, and eight (Eckroade et al. 1991). The state of Maryland's spending per child enrolled was \$1,754 in the 2002–03 school year (Barnett et al. 2003, 82).

In addition to the thirty-three program evaluations reviewed by Gilliam and Zigler (2000), the current researcher found two documented longitudinal evaluations of long-term program impact: one was conducted by the state of Illinois (Illinois State Board of Education 2001) and the other by the state of Delaware (Gamel-McCormick and Amsden 2002).

The Illinois evaluation followed a randomly selected sample of students who had previously participated in prekindergarten programs serving children living in poverty. It tracked the academic performance of these students from kindergarten to later grades. Each student's academic performance rating was determined by ISAT (Illinois Standards Achievement Tests) scores in conjunction with teacher-assigned academic performance rankings of “above average,” “average,” “below average,” and “deficient.” The teacher rankings were based on subjective judgments influenced by locally defined performance standards and assessment practices. No control group was formed. This study reports that in the Illinois downstate, the majority of students who had been in a prekindergarten program continued to do well: in the eighth grade, 72 percent of these students were ranked by their teachers as “average” or “above average” in reading; 77 percent were ranked as “average” or “above average” in mathematics and language (Illinois State Board of Education 2001, 13). Illinois Standards Achievement Tests (ISAT) results for downstate previous prekindergarten students in the eighth grade revealed that 41 percent were at the “meet” or “exceed” the state-standards level in mathematics; 64 percent were at that level in writing; and 72 percent were at that level in science (iii). In Chicago, 55 percent of the program students were ranked as “above average” or “average” in reading and mathematics in the eighth grade. In language, however, only about one third of the students were ranked as “above average” or “average” (20). The per pupil expenditure for the Illinois preschool program was \$3,094 in the 2002–03 school year (Barnett et al. 2003, 70).

The Illinois study used descriptive methods but was handicapped by a very high attrition rate. Of the 8,495 original subjects in kindergarten, only 649 subjects (7.6 percent) remained at the eighth grade in the downstate area (Illinois State Board of Education 2001, 15). No attrition analysis was conducted to determine if the subjects who stayed in the study were representative of the

original sample. Special caution should be taken when interpreting results from longitudinal studies that involve high attrition, lack a comparison group, use no statistical tests, and are based on local subjective judgments.

The Delaware study followed participants who were living in poverty and were served by the Early Childhood Assistance Program (ECAP) from the age of four to the third grade. Stratified sampling procedures were used for the comparison of students who received early intervention services and similar students who did not receive early intervention. These comparisons were conducted using statistical analyses such as ANOVA or means testing. A total of 70.6 percent of the original students remained in the study to the third grade—an attrition rate comparable to that seen in other similar longitudinal studies. Forty-two students who received ECAP or Head Start services at age four in 1996–97 remained in the study at the end of school year 2001–02. These students were compared with 109 students who were living in poverty at the time they began kindergarten but who had not received ECAP or Head Start services. The findings indicated that students who participated in ECAP or Head Start were significantly more likely to meet or exceed the standard in third-grade reading and mathematics as measured by the Delaware State Testing Program than those students living in poverty who did not receive ECAP or Head Start services. This study also found that students who received ECAP or Head Start services had significantly higher grades than their peers and a grade retention rate of less than half that of a comparable group of students (Gamel-McCormick and Amsden 2002, 12). The state spent \$5,287 for each child enrolled in ECAP during the 2002–03 school year (Barnett et al. 2003, 64).

Literature on the evaluation of large-scale, state-funded programs serving children at risk for school failure reveals that the studies of state-funded preschool programs vary considerably in their domains of interest, evaluation methodologies, and findings. But the findings are rather consistent that state-funded programs may help at-risk children enter school with a higher level of developmental competence and perform better in school during the critical early grades when compared to similar at-risk students with no preschool experience.

Studies of South Carolina early childhood development programs serving children at risk:

Act 135 requires the SDE to collect and analyze longitudinal data to determine the effects of child development programs on the later academic achievement of program participants by tracking the children through kindergarten and the first three years of elementary school and by examining their academic performance on appropriate performance measures. As of March 2003, seven statewide longitudinal studies on the impact of child development programs on the later academic achievements of program participants were available (SDE 1986, 1988, 1990, 1998, 1999, 2000, 2002).

Each study compared the program participants to a control group of randomly selected students who had not participated in the child development program but were comparable to the participants in their eligibility for the free- or reduced-price school lunch program. Three of these studies (1988, 1990, and 2002) followed participants beyond the first grade to second and third grades. The other four longitudinal studies examined participants' school readiness at grade one. All seven studies used inferential statistics to determine performance differences between participants and nonparticipants. Five of the seven evaluations revealed that at grades one, two, or three, program participants at risk for school failure had test scores significantly higher than their comparable peers who were not participants. In the other studies, program participants may

have had higher test scores than nonparticipants, but the difference was not significant. The 2002 study found that children who were at risk for school failure and had been served by child development programs at age four scored significantly higher than their comparable nonparticipant peers on school readiness as measured by the CSAB at grade one and on ELA and mathematics as measured by the PACT at grade three (see appendix B). Since 1984 the program has been supported by the state of South Carolina with EIA funds. The latest expenditure was \$1,467 per participant in the 2002–03 school year (Barnett et al. 2003, 120).

The Purpose of the Study

The purpose of this study is to determine the academic performance of four-year-old child development program participants when they were in grades four and five in school years 2000–01 and 2001–02. These children were identified as being at risk for school failure at age four and were served by child development programs in school year 1995–96.

Study Design and Methodology

Since students could not be randomly assigned to a treatment group, a quasi-experimental design was utilized in which a comparison group was established. A longitudinal match on the participants from the time that they were in the four-year-old child development program through their fifth-grade year was completed. The comparison group was constructed by randomly selecting children who were not participants in the four-year-old child development program but who were comparable to the participants in essential characteristics.

Population and Sample

All children who participated in child development programs in 1995–96 and qualified for free- or reduced-price school lunch at their enrollment in grade one (5,217 out of 9,977 valid records) were followed longitudinally through the fifth grade. The children chosen for the child development program were deliberately identified and recruited through a screening process utilizing the DIAL-R along with supplementary information about the child’s family such as education and income level. Most program participants were served in the half-day school classroom setting; others were served in a locally funded, full-day, center-based environment provided by some districts. Sixty-seven percent of the original subjects remained in the study at grade five. The purpose of selecting children not only at risk but also living in poverty was to ensure a group of participants as similar as possible to the comparison group that was randomly selected from nonparticipants who also were eligible for free- or reduced-price school lunch when they enrolled in the first grade. This was the same comparison group of 7,889 students who are the focus of section 1 of this document (“A Longitudinal Study of Later Academic Achievement, 1995–96 through 1999–2000,” above). Performance comparisons between child development program participants and nonparticipants were based on their PACT scores during the 2000–01 and 2001–02 school years, when they were in the fourth and fifth grades. Close to 60 percent (58.4 percent) of the original subjects selected were matched to the fifth-grade PACT test data.

The attrition in this longitudinal match was related to the lack of unique identification in the data collected by child development programs. This study saw an attrition rate comparable to that seen in similar studies, which ranged from less than 10 percent to about 25 percent per year (Gilliam and Zigler 2000). Table 10 reports the demographic features of the original program participants and those remaining in the matched data file in the fourth and fifth grades.

TABLE 10
Demographic Distributions of the Sample, by Percentage

		Program Participants Eligible for F/R Lunch at Grade 1			Nonprogram Participants Eligible for F/R lunch at Grade 1 Randomly Selected		
		Original Subjects N=5,217	Subjects Remaining at Grade 4 N=3,492	Grade 5 N=3,486	Original Subjects N=7,889	Subjects Remaining at Grade 4 N=4,640	Grade 5 N=4,618
Gender	Male	50.6%	47.0%	46.9%	52.3%	49.6%	49.5%
	Female	49.4%	53.0%	53.1%	47.7%	50.4%	50.5%
Race	African American	70.8%	69.9%	69.9%	66.4%	68.2%	68.2%
	Asian	0.5%	0.2%	0.2%	0.6%	0.4%	0.4%
	Caucasian	27.5%	29.0%	29.0%	31.1%	30.4%	30.4%
	Hispanic	1.0%	0.7%	0.7%	1.9%	1.0%	1.0%
	Other	0.2%	0.2%	0.2%	0.0%	0.0%	0.0%
Lunch Status	F/R eligible	100%	84.4%	83.3%	100.0%	85.2%	84.0%
	Not F/R eligible*	0.0%	15.6%	16.7%	0.0%	14.8%	16.0%

* Students' lunch status may have changed by the fourth or fifth grade.

Students in the two groups were comparable in ethnicity, gender, and family economic status. However, the comparison group consisted of students who, on the whole, had lesser degrees of risk for school failure than the child development program participants. The initial selection of the program and the comparison groups when their lunch status information was first available at first grade attempted to methodologically control the possible bias introduced by family economic status.

Data Utilized

Four data sources were utilized to generate the matched file in the investigation:

- survey data of early childhood development programs for four-year-olds in the 1995–1996 school year;
- statewide student information files from the 1997–98 school year;

- student performance data on the PACT in spring 2001, when the cohort group was in grade four; and
- student performance data on the PACT in spring 2002, when the cohort group was in grade five.

Data Analyses

Using the PACT in ELA and mathematics at the fourth and fifth grades, the SDE researcher compared the academic performance of program participants who were eligible for free-or reduced-price lunch with the performance of those nonparticipants who also were eligible for free-or reduced-price lunch. T-tests were used to determine the PACT mean scaled score differences between participants and nonparticipants. Analyses of variances (ANOVA) were applied for multiple comparisons of mean scaled score difference among three or more groups when subpopulations were examined. Chi-square tests were performed to determine if there were differences between participants and nonparticipants on the proportion of students meeting or exceeding the “proficient” standard on the PACT set by the state. The level of statistical significance was set at a probability value of .05 as the threshold; a probability below this threshold ($P < .05$) indicates that a difference of this magnitude could happen by chance less than 5 percent of the time.

Limitations to the Study

When one is designing educational program evaluation studies, certain limitations are inherently imposed. Limitations that need particular attention include the following:

- Due to ethical and practical considerations, individuals were not randomly assigned to treatment groups. With this limitation in mind, it is obvious that “true” experiments cannot be conducted when programs are being evaluated. This study employs quasi-experimental design; therefore, it is not feasible to completely rule out alternative explanations for the results.
- Uniform criteria for program implementation, instructional methods, and the quality of teachers’ professional development activities were not mandated at the time the data for this cohort group were collected.
- Children’s academic development and performance is a complex process that is influenced by factors other than preschool education. These factors include experiences at home, the school environment and the quality of teaching, and the characteristics of individual children. Data concerning the range of these factors were not available at the time this study was conducted.
- The comparison group for this study was randomly selected from nonprogram students eligible for the free- or reduced-price school lunch program. Though students in the treatment group were selected by eligibility for free- or reduced-price lunches, children enrolled in the four-year-old child development programs typically have significant readiness-deficiency

indicators other than low family income. Districts were required to identify and serve those students who were at the greatest risk for early school failure. Any selected comparison group likely would consist of students who, on the whole, have lesser degrees of risk.

- When interpreting the results for generality, the issue of attrition commonly associated with longitudinal data matching needs to be considered. The attrition rate in similar studies ranged from less than 10 percent to about 25 percent per year (Gilliam and Zigler 2000, 7–8). In this SDE study, the attrition rate was 38.2 percent across six years.
- There may be effects produced by additional educational services that the child development program participants may have received in the years from kindergarten through fifth grade. These effects were beyond the control of the current study.

For these reasons, statistical findings should be considered good but not exact. To maximize the internal and external validities, a comprehensive evaluation of the effectiveness and impact of child development programs for four-year-olds would require resolution of the above-described flaws in design and implementation.

Research Questions

1. What were the differences between the PACT mean scores of child development program participants and those of nonparticipants at the fourth and fifth grades? Did the mean scores differ by demographic characteristics? Did more program participants score at “proficient” or above as measured by the PACT?
2. Among child development program participants, which subgroup of students scored significantly higher on the PACT at the fourth and fifth grades?
3. What were the performance differences on the PACT between participants in half-day programs and those in full-day programs at the fourth and fifth grades?
4. Did program participation narrow the performance gap on the PACT between Caucasian and non-Caucasian students at the fourth and fifth grades?
5. What were the differences between the proportion of participants who took off-grade-level tests and the proportion of nonparticipants who took such tests at the fourth and fifth grades?

Findings

1. What were the differences between the PACT mean scores of child development program participants and those of nonparticipants at the fourth and fifth grades? Did the mean scores differ by demographic characteristics? Did more program participants score at “proficient” or above as measured by the PACT?

Program participants who were eligible for the free- or reduced-price lunch program at grade one were tracked through the fourth and fifth grades. The mean scaled scores of program participants on the PACT ELA and mathematics tests were compared with those of the nonparticipants. The results demonstrated significant performance differences in both ELA and mathematics at grades four and five related to program participation: children who were at risk of school failure and who participated in child development programs scored significantly higher on the PACT at grades four and five than their peers. Figures 3 and 4 show the details.

Figure 3

Mean Score Comparison between Child Development Program Participants and Nonparticipants: Fourth-Grade PACT Performance Spring 2001

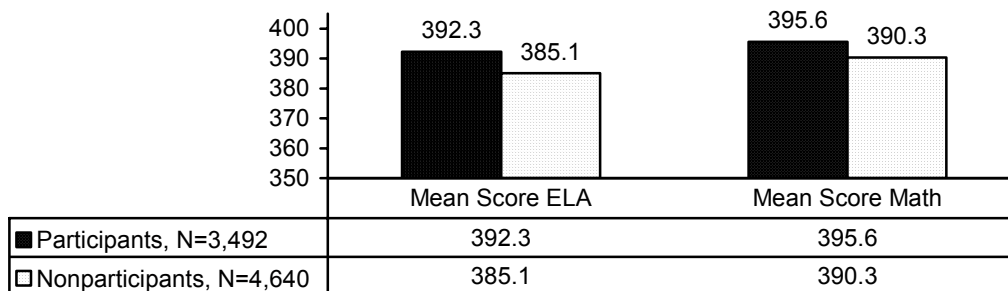
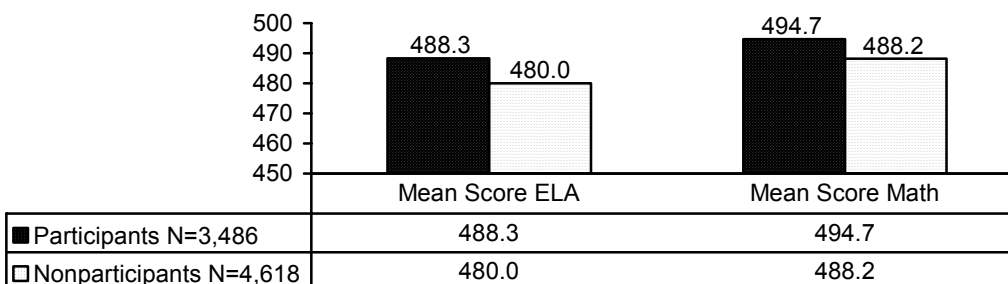


Figure 4

Mean Score Comparison between Child Development Program Participants and Nonparticipants: Fifth-Grade PACT Performance Spring 2002



In a comparison of PACT scores between participants and nonparticipants, program participants scored significantly higher on state assessments on both ELA and mathematics at grade four than their nonparticipant peers ($P < .05$). When fifth-grade test scores were compared, findings again significantly favored program participants on both subject areas ($P < .05$).

By Demographic Category

Further analyses were conducted to compare the performance of participants and nonparticipants by different demographic characteristics. Mean scaled scores on the PACT for grades four and five were used to examine student performance by gender, race, and lunch status.

Findings showed that all subpopulations of program participants significantly outscored nonparticipants. Tables 11 and 12 provide the details.

TABLE 11

**Mean Score Comparison between Child Development Program
Participants and Nonparticipants: Fourth-Grade Performance on the PACT
Spring 2001, by Demographic Category**

Demographic Category	Mean Scores ELA		Mean Scores Math		Probability Value	
	Program N=3,492	Nonprogram N=4,640	Program N=3,492	Nonprogram N=4,640	ELA	Math
Male	387.2	377.5	394.4	386.5	< .05	< .05
Female	396.8	395.8	396.7	394.1	< .05	< .05
Caucasian	395.7	390.7	399.1	397.1	< .05	< .05
Non-Caucasian	390.9	382.7	394.2	387.4	< .05	< .05
Eligible for free lunch program	390.1	381.1	393.5	387.0	< .05	< .05
Eligible for reduced-price lunch program	396.1	395.3	398.9	396.7	< .05	< .05
Not eligible for the subsidized lunch program	399.7	397.7	403.2	402.3	< .05	< .05

TABLE 12

**Mean Score Comparison between Child Development Program
Participants and Nonparticipants: Fifth-Grade Performance on the PACT
Spring 2002, by Demographic Category**

Demographic Category	Mean Scores ELA		Mean Scores Math		Probability Value	
	Program N=3,486	Nonprogram N=4,618	Program N=3,486	Nonprogram N=4,618	ELA	Math
Male	482.8	470.3	492.1	482.2	< .05	< .05
Female	493.0	489.5	496.9	494.1	< .05	< .05
Caucasian	492.9	487.5	500.7	498.0	< .05	< .05
Non-Caucasian	486.4	476.7	492.2	483.9	< .05	< .05
Eligible for free lunch program	484.3	475.9	490.7	484.4	< .05	< .05
Eligible for reduced-price lunch program	494.9	486.7	501.6	493.8	< .05	< .05
Not eligible for the subsidized lunch program	500.2	494.1	506.4	501.8	< .05	< .05

Performance Levels

The state assessment system classifies students' PACT performance into four achievement levels: "below basic," "basic," "proficient," and "advanced." The percentage of participants meeting or exceeding the "proficient" level on the PACT was compared to the percentage of nonparticipants meeting or exceeding the "proficient" level at grades four and five. Table 13 gives the details.

TABLE 13

**Percentage of Students Scoring at or above Proficient on the PACT
Program Participants Compared to Nonparticipants
Grades Four and Five, Spring 2001 and 2002**

	Participants % Meeting or Exceeding Proficient		Nonparticipants % Meeting or Exceeding Proficient		Probability Value	
	ELA	Math	ELA	Math	ELA	Math
Grade 4	23.9%	15.6%	23.2%	15.2%		
Grade 5	13.3%	17.7%	12.3%	15.6%		< .05

A significantly larger proportion of participants than nonparticipants met or exceeded “proficient” on mathematics at grade five. Findings favored program participants in fourth- and fifth-grade ELA and fourth-grade mathematics. However, the differences were not large enough to be statistically significant.

2. Among child development program participants, which subgroup of students scored significantly higher on the PACT at the fourth and fifth grades?

PACT scores for child development program participants were compared by gender, race, and lunch status at grades four and five. Lunch status for 16.7 percent of the program participants changed from eligible for the subsidized lunch program to ineligible between the first and the fifth grade. Ethnic groups of participating children with fewer than thirty students were reclassified into an “other” racial group to yield valid statistical results. Female participants consistently outscored male participants from grade four to grade five. Participating children not eligible for the free lunch program scored significantly higher than children who were eligible for either the free or the reduced-price lunch program. Mean scores of African American participants were lower than those of Caucasian participants and of participants from other racial groups. Tables 14 through 16 display the comparisons by gender, race, and lunch status, respectively.

TABLE 14
PACT Mean Score Comparison within Child Development Programs
Spring 2001 and 2002, by Gender

Gender	Grade 4		Grade 5	
	PACT Mean Scores		PACT Mean Scores	
	Statewide		Statewide	
	ELA N=3,492	Math N=3,492	ELA N=3,486	Math N=3,486
Female (F)	396.8	396.7	493.1	497.0
Male (M)	387.2	394.4	482.2	492.1
Gender group comparison results	F > M	F > M	F > M	F > M
Probability value	< .05	< .05	< .05	< .05

TABLE 15
PACT Mean Score Comparison within Child Development Programs
Spring 2001 and 2002, by Race

Race	Grade 4 PACT Mean Scores		Grade 5 PACT Mean Scores	
	Statewide		Statewide	
	ELA N=3,492	Math N=3,492	ELA N=3,486	Math N=3,486
Caucasian	395.7	399.1	492.9	500.7
African American (AA)	390.8	394.1	486.1	491.9
Other	395.8	398.9	504.0	509.7
Racial group comparison results	Insignificant	Insignificant	Caucasian & Other > AA	Caucasian & Other > AA
Probability value			< .05	< .05

TABLE 16
PACT Mean Score Comparison within Child Development Programs
Spring 2001 and 2002, by Lunch Status

Lunch Status	Grade 4 PACT Mean Scores		Grade 5 PACT Mean Scores	
	Statewide		Statewide	
	ELA N=3,492	Math N=3,492	ELA N=3,486	Math N=3,486
Not eligible for free- or reduced-price (NF/R) lunch	399.7	403.2	500.2	506.4
Eligible for reduced-price lunch (R)	396.1	398.9	494.9	501.6
Eligible for free lunch (F)	390.1	393.5	484.3	490.7
Lunch group comparison results	NF/R & R lunch > F lunch	NF/R & R lunch > F lunch	NF/R & R lunch > F lunch	NF/R & R lunch > F lunch
Probability value	< .05	< .05	< .05	< .05

Analyses on the performance of participants within program by subgroups showed that female participants consistently scored higher than male participants on both ELA and mathematics as measured by the PACT in both grades four and five and that the economic status of participants was positively associated with participants' academic performance. Participants not eligible for the free or reduced-price lunch program scored significantly higher on both ELA and mathematics at grades four and five than the participants who were eligible for the free lunch

program. Ethnic origin was found to be significantly associated with participants' performance only at grade five. Caucasians and thirty-nine Asian and Hispanic participants combined into an "other" category scored significantly higher than African Americans on both fifth-grade ELA and mathematics.

3. What were the performance differences on the PACT between participants in half-day programs and those in full-day programs at the fourth and fifth grades?

The 455 participants who were served at center-based, full-day programs were compared to 3,034 participants from half-day programs located in school settings. Participants' PACT mean scores were analyzed for both ELA and mathematics at grades four and five.

Findings indicated no significant difference in student performance as measured by the PACT between participants in full-day and those in half-day programs, though the mean scores favored participants from half-day programs on fourth-grade ELA and fifth-grade ELA and mathematics. However, readers should be cautious in interpreting results based on only 455 students.

4. Did program participation narrow the performance gap on the PACT between Caucasian and non-Caucasian students at the fourth and fifth grades?

The gap between the mean scores of Caucasians and non-Caucasians on the PACT ELA and mathematics assessments were examined for both program participants and nonparticipants at the fourth and fifth grades.

Program participation significantly helped children at risk of academic deficiency to close the academic achievement gap between Caucasians and non-Caucasians on the PACT at the fourth and fifth grades. Table 17 and figures 5 through 8 depict the differences.

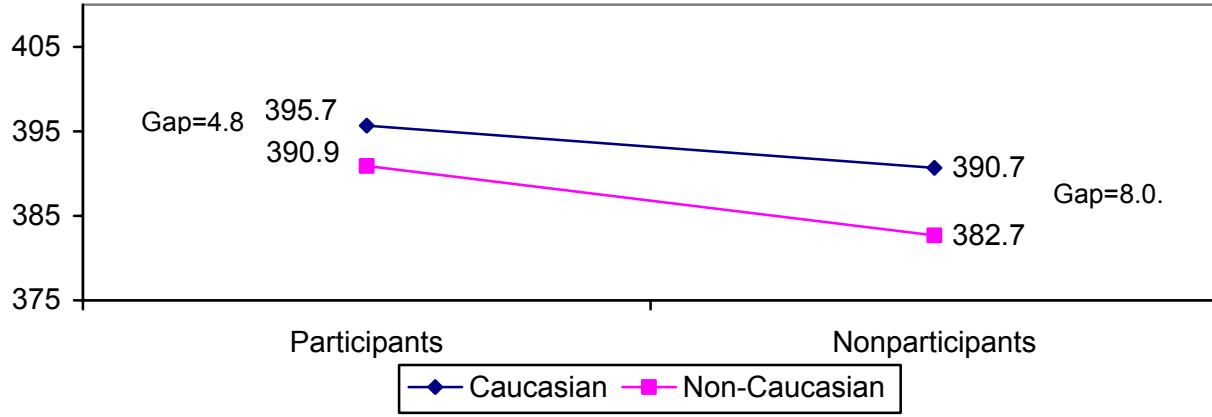
TABLE 17

**PACT Mean Score Gaps between Caucasian and Non-Caucasian Students:
Program Participants Compared to Nonparticipants, Spring 2001 and 2002**

	Participants Mean Score Gap between Caucasians and Non-Caucasians				Nonparticipants Mean Score Gap between Caucasians and Non-Caucasians			
	ELA		Math	Probability Value	ELA	Probability Value	Math	Probability Value
Grade 4	4.8	Insignificant	4.9	Insignificant	8.0	< .05	9.7	< .05
Grade 5	6.5	Insignificant	8.5	< .05	10.8	< .05	14.1	< .05

Figure 5

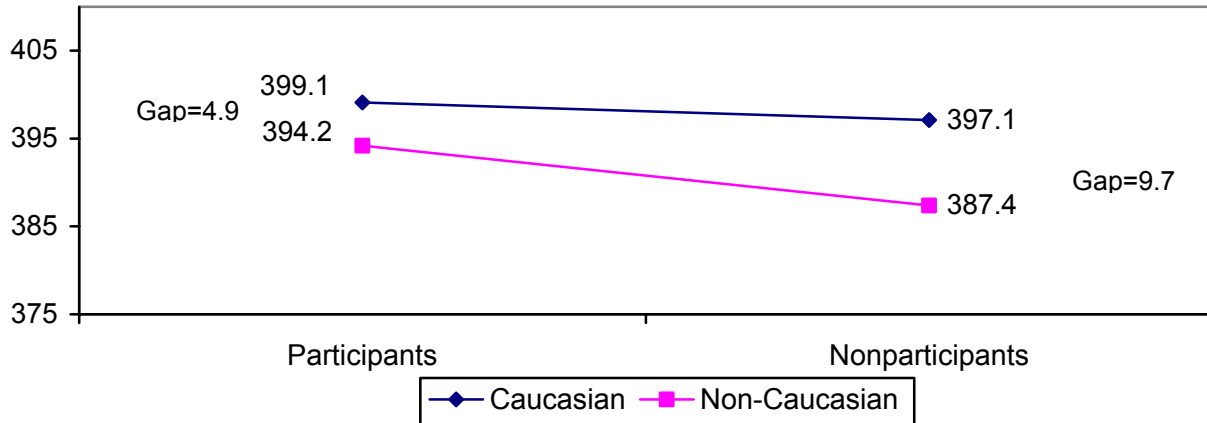
**PACT Mean Score Gaps between Caucasian and Non-Caucasian Students:
Program Participants Compared to Nonparticipants
Grade Four, English Language Arts, Spring 2001**



Standard error of measurement = 3.8 (Source: SDE Office of Assessment)

Figure 6

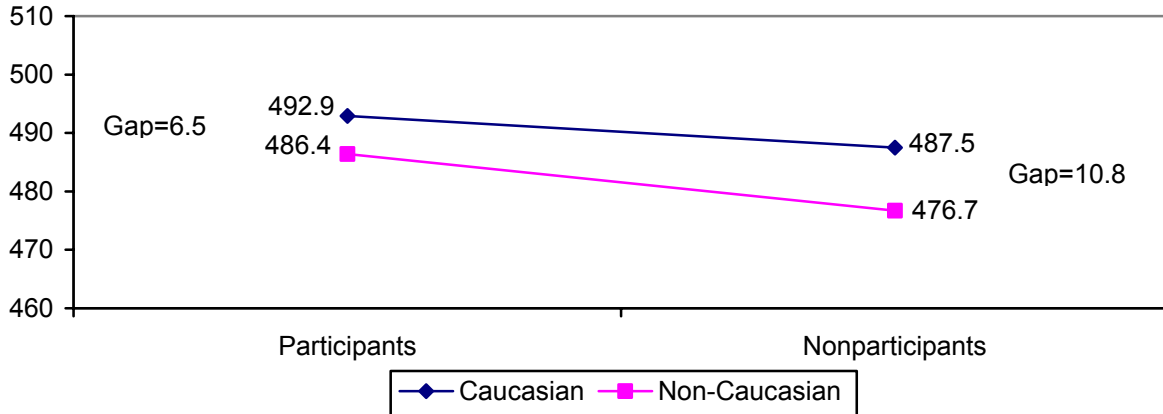
**PACT Mean Score Gaps between Caucasian and Non-Caucasian Students:
Program Participants Compared to Nonparticipants
Grade Four, Mathematics, Spring 2001**



Standard error of measurement = 5.9 (Source: SDE Office of Assessment)

Figure 7

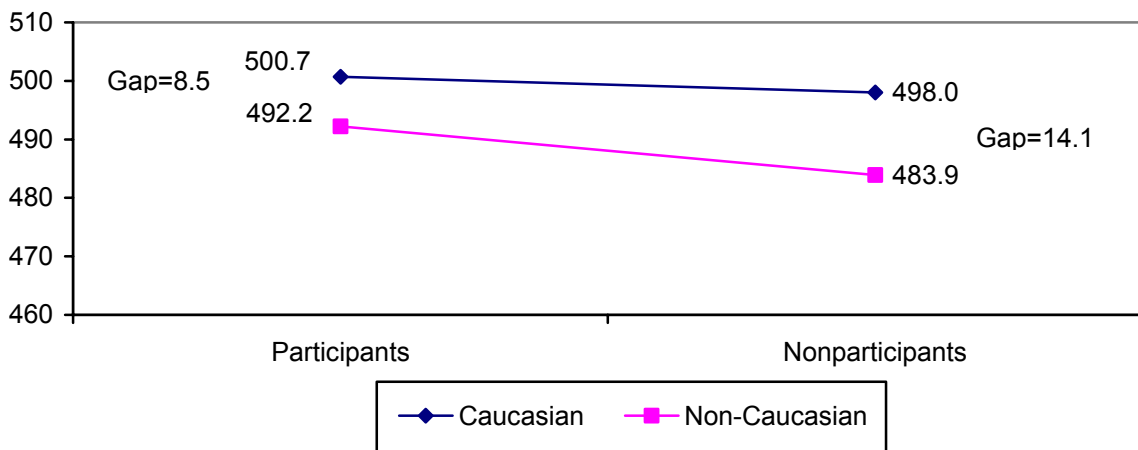
**PACT Mean Score Gaps between Caucasian and Non-Caucasian Students:
Program Participants Compared to Nonparticipants
Grade Five, English Language Arts, Spring 2002**



Standard error of measurement = 4.1 (Source: SDE Office of Assessment)

Figure 8

**PACT Mean Score Gaps between Caucasian and Non-Caucasian Students:
Program Participants Compared to Nonparticipants
Grade Five, Mathematics, Spring 2002**



Standard error of measurement = 5.7 (Source: SDE Office of Assessment)

Analyses indicated that program participation helped in closing the performance gap between Caucasian and non-Caucasian students. The performance gaps between the two racial categories for the child development participants were much narrower than that of nonparticipants and was statistically insignificant in fourth-grade ELA and mathematics and fifth-grade ELA as measured by the PACT. Compared to program participants, the performance gaps between Caucasian and non-Caucasian nonparticipants on the PACT were wider and statistically significant in ELA and mathematics for both grades four and five.

5. What were the differences between the proportion of participants who took off-grade-level tests and the proportion of nonparticipants who took such tests at the fourth and fifth grades?

The proportion of participants who took off-grade-level tests (i.e., students taking a grade-level test from a grade level lower than their current grade level) on the PACT was compared with that of nonparticipants. A significantly smaller proportion of child development program participants took off-grade-level tests on the PACT compared to their nonparticipant peers.

These findings show that participation in a child development program significantly helped reduce the number of students needing to take off-grade-level tests in grade four and grade five. Table 18 shows the differences in detail.

TABLE 18
Percentage of Students Taking Off-Grade-Level Tests:
Program Participants Compared to Nonparticipants
PACT, Spring 2000 and 2001

	Off-Grade-Level Tests ELA		Off-Grade-Level Tests Mathematics		Probability Value	
	Participants	Nonparticipants	Participants	Nonparticipants	ELA	Math
Grade 4	3.5%	6.1%	2.8%	4.7%	< .05	< .05
Grade 5	4.2%	7.7%	3.5%	6.5%	< .05	< .05

Implications and Conclusions

Many research studies have provided adequate proof that high-quality preschool programs can produce positive effects for children at risk of school failure. However, the evaluation, research methodology, and findings on large-scale, state-funded programs varied widely. With the limitations that were beyond the control by the current researcher, the findings from this study should be considered good but not exact. This study provides consistent statewide evidence that the South Carolina state-funded early childhood program has a long-term positive impact on children's academic performance. It suggests that the impact was sustained six years after the participants had left the program.

This study focused on the academic performance of fourth- and fifth-grade students who had participated in South Carolina child development programs in school year 1995–96 at age four. The current study traced the cohort group studied in 2002 to the fourth and fifth grades. It compared the academic achievement of participants and nonparticipants as measured by the PACT in the spring of 2001 and 2002. Test scores of program participants were examined by demographic characteristics and program type. Analyses were carried further on the performance gap between Caucasian students and non-Caucasian students. Lastly, this study examined whether program participation could help reduce the number of students needing to take off-grade-level tests.

The findings of this study lead to the following conclusions:

- Participation in child development programs for four-year-olds had a positive long-term impact on participants' academic performance six years after the students had left the program when compared to randomly selected similar students who did not participate in a child development program. By definition, program participants were identified as at risk for school failure based not only on low family income but also on other factors such as low DIAL-R scores, low maternal education level, a home language other than English, and health problems. In spite of their risk levels, the program participants significantly outscored nonparticipants on both ELA and mathematics as measured by the PACT at grades four and five. All subgroups of participants scored significantly higher than nonparticipants when comparisons were conducted on the population disaggregated by gender, race, and lunch status. The findings suggest that the program impact on children's academic performance could continue to grade five.
- Test performance by program participants differed among demographic groups. Female participants scored significantly higher than their male peers in ELA and mathematics at both grades four and five. Fifth-grade Caucasian and Asian and Hispanic participants combined into an "other" category both significantly outscored African American participants on ELA and mathematics. Family income was positively related to students' academic performance. At both grades four and five, program participants eligible for the free lunch program scored significantly lower on the PACT ELA and mathematics than those participants not eligible for the subsidized lunch program and those participants eligible for reduced-price lunch. This finding supported the results from the 2002 study that among program participants, students' later academic performance varied by demographic group.
- No performance differences were shown by the test scores of fourth- and fifth-grade students served by full-day programs and those served by half-day programs. This finding again was consistent with the results from the previous study of 2002, in which student performance at the third grade was similar for participants whether they were served by full-day or half-day programs.
- Program participation appeared helpful in significantly narrowing the performance gap between Caucasian and non-Caucasian participants. The performance gap between Caucasian and non-Caucasian students not participating in a child development program was greater than that for the program participants.
- Program participation appears to significantly reduce the percentage of students taking off-grade-level tests. This finding may indicate that early intervention allows at-risk students to keep up with their classmates.

Recommendations

The following recommendations are offered to help researchers and policy makers conduct further studies to understand details of how program design and program implementation may influence the fulfillment of program goals:

- Evaluations built on the knowledge base of educational interventions that have been proven effective through randomized controlled trials should be conducted. The experiment could be managed not only in a small demonstration project but also in districts where the most needy children or communities are located. Both quantitative and qualitative data should be used to offer empirical proof of model practices that best fulfill program goals.
- Studies based on multiple data sources should be conducted for program improvement. Efforts should be made to obtain information on classroom environment and teacher qualifications and to extract specific indicators that are closely related to the intended outcome of the program.
- A study using national achievement data, such as those from the National Assessment of Educational Progress (NAEP), should provide further evidence for policy analysis.

APPENDIX A

Districts Providing MAT7 Test Data, Spring 2002

Lowcountry	Midlands	Upstate
Beaufort	Aiken	Cherokee
Charleston	Barnwell 45	Chester
Georgetown	Kershaw	Greenville
Horry	Lexington 1	Greenwood 50
Marion 3	Lexington 3	Laurens 55
Orangeburg 3	Lexington 4	Oconee
	Richland 2	Spartanburg 1
		Spartanburg 4
		Spartanburg 7
		York 3

APPENDIX B

**Summary of Longitudinal Studies on South Carolina
Early Childhood Development Programs for Four-Year Olds, 1986–2004**

Year Reported and Document Title*	SDE Investigator	Cohort Group Studied	Comparison Group	Data Analyses Methods	Outcome Measured	Result
1986 <i>Early Childhood Development Programs: Half-Day Programs for Four-Year Olds. 1983–84 School Year: Large Sample Evaluation Report</i>	M. Nadir Atash	1983–84 school year	Random sampling of nonparticipants eligible for free lunch	Descriptive and inferential statistics (MANOVA [multiple analysis of variance], t-test)	1. First-grade school readiness as measured by the CSAB 2. First-grade academic performance on the BSAP	1. At grade one, the mean scores on the CSAB were similar between participants and nonparticipants. 2. At grade one, participants scored significantly higher than nonparticipants on BSAP reading and math.
1988 <i>The Longitudinal Evaluation of Early Childhood Education: Interim Report. The Half-Day Child Development Program for Four-Year-Olds</i>	Molly M. Jones	1983–84 school year	Randomly selected nonparticipants	Descriptive and inferential statistics (chi-square, ANOVA)	1. First-grade school readiness as measured by the CSAB 2. Academic performance on the BSAP at grades one and two	1. At grade one, the mean scores on the CSAB were similar between participants and nonparticipants. 2. At grades one and two, student BSAP performances were similar between participants and nonparticipants.
1990 <i>The Longitudinal Evaluation of Early Childhood Education: Third Interim Report-Revised. The Half-Day Child Development Program for Four-Year Olds</i>	Molly M. Jones	1985–86 school year	Randomly selected nonparticipants	Descriptive and inferential statistics (chi-square, t-test)	1. First-grade school readiness as measured by the CSAB 2. Academic performance on the BSAP at grades one, two, and three	1. The mean school readiness scores on the CSAB were similar between participants and nonparticipants. 2. At grades one, two, and three, student performances on the BSAP reading test were similar between participants and nonparticipants. Student performances on the BSAP math test were similar between participants and nonparticipants at grade one, but in both grades two and three, nonparticipants scored significantly higher than nonparticipants.

**Summary of Longitudinal Studies on South Carolina
Early Childhood Development Programs for Four-Year Olds, 1986–2004**

Year Reported and Document Title*	SDE Investigator	Cohort Group Studied	Comparison Group	Data Analyses Methods	Outcome Measured	Result
1998 <i>The Longitudinal Research Report on the Early Childhood Development Program: The Half-Day Child Development Program for Four-Year-Olds, 1995–96</i>	Wei Yao	1995–96 school year	Random selection of nonparticipants eligible for free lunch	Descriptive and inferential statistics (chi-square)	First-grade school readiness as measured by the CSAB	The percentage of students ready for school at grade one was similar for participants and nonparticipants.
1999 <i>A Longitudinal Research Report on the Early Childhood Development Program: The Half-Day Child Development Program for Four-Year-Olds, 1996–97</i>	Wei Yao	1996–97 school year	Random selection of nonparticipants eligible for free lunch	Descriptive and inferential statistics (chi-square)	First-grade school readiness as measured by the CSAB	The percentage of students ready for school at grade one was similar for participants and nonparticipants.
2000 <i>A Longitudinal Research Report on the Early Childhood Development Program: The Half-Day Child Development Program for Four-Year-Olds, 1997–98.</i>	Wei Yao	1997–98 school year	Random selection of nonparticipants eligible for free lunch	Descriptive and inferential statistics (chi-square)	First-grade school readiness as measured by the CSAB	The percentage of students ready for school at grade one was similar for participants and nonparticipants.

**Summary of Longitudinal Studies on South Carolina
Early Childhood Development Programs for Four-Year Olds, 1986–2004**

Year Reported and Document Title*	SDE Investigator	Cohort Group Studied	Comparison Group	Data Analyses Methods	Outcome Measured	Result
2002 <i>Child Development Programs for Four-Year-Olds: Student and Program Characteristics, Longitudinal Study of Academic Achievement, and Current Parent Perceptions</i>	Wei Yao	1995–96 school year	Random selection of nonparticipants eligible for free- or reduced-price lunch	Inferential statistics (ANOVA, ANCOVA)	<ol style="list-style-type: none"> 1. First-grade school readiness as measured by the CSAB 2. Performance on the MAT7 at grade two 3. Performance on the PACT at grade three 	<ol style="list-style-type: none"> 1. At grade one, participants scored significantly higher on the CSAB than their comparable peers. 2. At grade two, participants scored higher on the MAT7 reading and math tests, but the difference was not large enough to be statistically significant. 3. At grade three, participants scored significantly higher on the PACT in both ELA and math than their comparable peers.

* Documents available at the South Carolina State Library

References

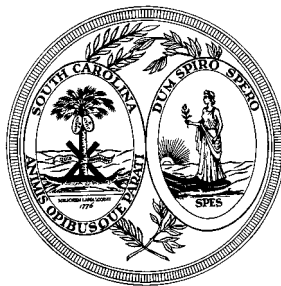
- Barnett, W. Steven. 1995. "Long-Term Effects of Early Childhood Programs on Cognitive and School Outcomes." *Future of Children* 5, no. 3:25–50. Also available online at http://www.futureofchildren.org/usr_doc/vol5no3ART2.pdf.
- . 2002. "Early Childhood Education." In *School Reform Proposals: The Research Evidence*, edited by Alex Molnar. Greenwich, CN: Information Age. Also available online at <http://www.asu.edu/educ/eps1/EPRU/documents/EPRU%202002-101/epru-2002-101.htm>.
- Barnett, W. Steven, et al. 1987. "The Efficacy of Public Preschool Programs and the Relationship of Program Quality to Efficacy." *Educational Evaluation and Policy Analysis* 10, no. 1:37–49.
- , et al. 2003. *The State of Preschool: 2003 State Preschool Yearbook*. National Institute of Early Education Research. <http://nieer.org/yearbook/> (accessed October 26, 2004).
- Campbell, Frances A., and Craig T. Ramey. 1999. *Early Learning, Later Success: The Abecedarian Study—Early Childhood Educational Intervention for Poor Children. Executive Summary*. Frank Porter Graham Child Development Center. <http://www.gomasa.org/Resources/EduPortal/EarlyChildhood/abecedarian.pdf> (accessed October 25, 2004).
- Campbell, Frances A., et al. 1998. "Early Childhood Programs and Success in School: The Abecedarian Study." In *Early Care and Education for Children in Poverty: Promises, Programs, and Long-Term Results*, edited by W. Steven Barnett and Sarane Spence Boocock. Albany: State University of New York Press.
- , et al. 2001. "The Development of Cognitive and Academic Abilities: Growth Curves from an Early Childhood Educational Experiment." *Developmental Psychology* 37, no. 2:231–42.
- "Carolina Abecedarian Project." 2003. Research Forum, National Center for Children in Poverty. http://www.researchforum.org/project_general_228.html (accessed November 5, 2004).
- Denton, Kristin, and Jerry West. 2002. *Children's Reading and Mathematics Achievements in Kindergarten and First Grade*. Washington, DC: U.S. Department of Education, National Center for Education Statistics. Also available online at <http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2002125>.
- Eckroade, Ginger, et al. 1991. *An Analysis of the Long-Term Effects of the Extended Elementary Education Prekindergarten Program*. Baltimore: Maryland State Department of Education.

- Gamel-McCormick, Michael, and Deborah Amsden. 2002. *Investing in Better Outcomes: The Delaware Early Childhood Longitudinal Study*. Newark: Center for Disabilities Studies, University of Delaware. Also available online at http://www.doe.state.de.us/early_childhood/research/DECLS02.pdf.
- Gilliam, Walter S., and Edward F. Zigler. 2000. "A Critical Meta-Analysis of All Impact Evaluations of State-Funded Preschool from 1997 to 1998: Implications for Policy, Service Delivery and Program Evaluation." *Early Childhood Research Quarterly* 15, no. 4 (winter): 441–73. Also available online at <http://nieer.org/resources/research/StatePreKMeta.pdf>.
- Gilliam, Walter S., and Carol H. Ripple. 2004. "What Can Be Learned from State-Funded Prekindergarten Initiatives: A Data-Based Approach to the Head Start Devolution Debate." In *The Head Start Debates*, edited by Edward Zigler and Sally J. Styfco. Baltimore, MD: Brookes.
- Illinois State Board of Education. Research Division. 2001. *Illinois Prekindergarten Program for Children at Risk of Academic Failure: FY 2000 Evaluation Report*. Chicago: Illinois State Board of Education. Also available online at <http://www.isbe.net/research/pdfs/atrisk.pdf>.
- NIEER. 2004. "A Benefit-Cost Analysis of the Abecedarian Early Childhood Intervention." National Institute for Early Education Research. <http://nieer.org/docs/index.php?DocID=57> (accessed November 1, 2004).
- Ramey, Craig T., and Sharon L. Ramey. 1998. "Early Intervention and Early Experience." *American Psychologist* 53:109–20.
- Reynolds, Arthur J. 2002. "Early Childhood Interventions: Knowledge, Practice and Policy." *Focus* 22, no. 1:112–17. Also available online at <http://www.ssc.wisc.edu/irp/focus/foc221-part4.pdf>.
- Reynolds, Arthur J., et al. 2001. "Long-Term Effects of an Early Childhood Intervention on Educational Achievement and Juvenile Arrest." *Journal of the American Medical Association* 285 (May) 18:2339–46.
- Schweinhart, Lawrence J. 2001. *Recent Evidence on Preschool Programs* (report no. EDO-PS-01-9). Champaign, IL: ERIC Clearinghouse on Elementary and Early Childhood Education. ERIC Document Reproduction Service no. ED 458 046. Also available online at <http://ceep.crc.uiuc.edu/ecearchive/digests/2001/schweinhart01.html>.
- . 2002. "How the How the High/Scope Perry Preschool Study Grew: A Researcher's Tale." *Phi Delta Kappa Research Bulletin*, no. 32 (June). Also available online at <http://www.pdkintl.org/edres/resbul32.htm>.
- . 2003. "Benefits, Costs, and Explanation of the High/Scope Perry Preschool Program." Paper presented at the meeting of the Society for Research in Child Development, Tampa, Florida, April 26. Available online at <http://www.highscope.org/Research/PerryProject/Perry-SRCD-2003.pdf>.

- Schweinhart, Lawrence J., Helen V. Barnes, and David P. Weikart. 1993. *Significant Benefits: The High/Scope Perry Preschool Study through Age 27*. Monographs of the High/Scope Educational Research Foundation, no. 10. Ypsilanti, MI: High/Scope Press.
- SDE. Office of Research. 1986. *Early Childhood Development Programs: Half-Day Programs for Four-Year Olds. 1983–84 School Year: Large Sample Evaluation Report*. Columbia: South Carolina Department of Education.
- . 1988. *The Longitudinal Evaluation of Early Childhood Education: Interim Report. The Half-Day Child Development Program for Four-Year-Olds*. Columbia: South Carolina Department of Education.
- . 1998. *The Longitudinal Research Report on the Early Childhood Development Program: The Half-Day Child Development Program for Four-Year-Olds, 1995–96*. Columbia: South Carolina Department of Education.
- . 1990. *The Longitudinal Evaluation of Early Childhood Education: Third Interim Report-Revised. The Half-Day Child Development Program for Four-Year-Olds*. Columbia: South Carolina Department of Education.
- . 1999. *A Longitudinal Research Report on the Early Childhood Development Program: The Half-Day Child Development Program for Four-Year-Olds, 1996–97*. Columbia: South Carolina Department of Education.
- . 2000. *A Longitudinal Research Report on the Early Childhood Development Program: The Half-Day Child Development Program for Four-Year-Olds, 1997–98*. Columbia: South Carolina Department of Education.
- . 2002. *Child Development Programs for Four-Year-Olds: Student and Program Characteristics, Longitudinal Study of Academic Achievement, and Current Parent Perceptions*. What Is the Penny Buying for South Carolina Series. Columbia: South Carolina State Board of Education.
- U.S. Department of Health and Human Services. 2003. *State-Funded Pre-Kindergarten: What the Evidence Shows*. Washington, DC: Office of the Assistant Secretary for Planning and Evaluation. Also available online at <http://aspe.hhs.gov/hsp/state-funded-pre-k/>.

Total Printing Costs \$84.75
Total Units Printed..... 75
Cost per Unit..... \$1.13

The South Carolina Department of Education does not discriminate on the basis of race, color, national origin, sex, or disability in admission to, treatment in, or employment in its programs and activities. Inquiries regarding the nondiscrimination policies should be made to the director of the Office of Human Resources, 1429 Senate Street, Columbia, South Carolina 29201, 803-734-8505.



A publication of the
State Board of Education

1429 Senate Street, Columbia, South Carolina 29201
Phone: 803-734-8492 Fax: 803-734-8624
<http://www.myschools.com>