

Chapter 2. *Bright Beginnings* and *Creative Curriculum*: Vanderbilt University (Tennessee site)

Curriculum

The Vanderbilt University (Tennessee) researchers evaluated *Bright Beginnings* and *Creative Curriculum*.¹

Bright Beginnings

Bright Beginnings is an integrated curriculum with a focus on language and early literacy. It is based in part on the *High/Scope* and *Creative Curriculum* models, with an added focus on literacy skills that are designed to promote school readiness. The curriculum goals are to provide a child-centered, literacy-focused program that is consistent with developmentally appropriate practice and to include instruction that addresses the cognitive, social, emotional, and physical development of young children. The curriculum was especially designed to provide continuity in the pre-kindergarten to second-grade curricula. *Bright Beginnings* includes nine curriculum units that are linked to the program components:

- language and literacy;
- mathematics;
- social and personal development;
- healthful living;
- scientific thinking;
- social studies;
- creative arts;
- physical development; and
- technology.

The classroom environment is designed to encourage children's active exploration and interaction with adults, other children, and classroom materials. Teachers conduct ongoing assessments of children as they engage in a range of classroom activities. The curriculum also includes a Family-School Connection link. Parents sign a parent-school partnership agreement that requires a parent/caregiver to be actively engaged in the child's education.

Creative Curriculum

Creative Curriculum is a comprehensive curriculum for 3- to 5-year-old children. The curriculum addresses four areas of development:

- social/emotional;
- physical;
- cognitive; and
- language development.

¹ The University of North Carolina at Charlotte research team also evaluated *Creative Curriculum*.

Creative Curriculum requires the physical space of the classroom to be structured into 10 interest areas: blocks, dramatic play, toys and games, art, library, discovery, sand and water, music and movement, cooking, and computers. Time is also allotted for outdoor activities. The 10 interest areas are designed to address curriculum content, such as literacy, mathematics, science, social studies, the arts, technology, and process skills, such as observing, exploring, and problem solving. *Creative Curriculum* includes a Developmental Checklist teachers are asked to use in ongoing assessments of child progress.

Sample

The Tennessee research team recruited 36 public pre-kindergarten classrooms in seven school districts in six different counties. All of the selected programs were full-day pre-kindergarten programs. Teachers were recruited in July of the preschool year, curriculum training occurred in August, and parental consent was obtained in late August 2003 and early September 2003. A total of 36 teachers/classrooms and 558 parents and children were recruited for participation in the site-level study. A subset of that sample (21 classrooms and teachers) including 309 children and 300 parents (103 in the *Bright Beginnings* treatment group, 101 in the *Creative Curriculum* treatment group, and 105 in the control group) were included in the PCER study sample. Data were collected on 309 children and 252 parents at the time of the fall baseline data collection.

In the follow-up year of the evaluation, the sample of schools went from 19 preschools to 64 schools in kindergarten. The evaluation sample of classrooms went from 21 preschool to 134 kindergarten classrooms. The kindergarten sample included 307 children and 298 parents from the original sample of participants. Data were collected on 300 children and 232 parents.

Children and Families

The children were 4.5 years old at the time of baseline data collection and slightly more than half (52%) were male. The sample of children was White (80%), African American (18%), and Hispanic (11%). Table 2.1 provides additional information on the demographic characteristics of the children in the study sample. At baseline, a higher percentage of control group parents reported that their child had an Individualized Education Plan (IEP) relative to those assigned to the *Bright Beginnings* and *Creative Curriculum* conditions (33% vs. 13% and 12%, $p < .01$).

Table 2.1. Child demographic characteristics for *Bright Beginnings* and *Creative Curriculum*

Characteristics	Full sample n = 309	Curriculum comparison		
		Control n = 105	Treatment 1 ¹ n = 101	Treatment 2 ² n = 103
Age at baseline (years), mean	4.5	4.5	4.6	4.5
Gender (% male)	52.1	48.6	53.5	54.4
Race/ethnicity (%)				
White, non-Hispanic	79.6	84.0	74.4	80.0
African American, non-Hispanic	6.5	‡	‡	‡
Hispanic	10.8	10.6	11.1	10.5
Asian or Pacific Islander	‡	‡	‡	‡
Native American	‡	‡	‡	‡
Multiple/other	‡	‡	‡	‡
Child disability status (parent reported, %)	18.7	32.5**	11.7	13.0

‡ Reporting standards not met. Values suppressed to protect participant confidentiality.

** $p < .01$

¹Treatment 1 is *Creative Curriculum*.

²Treatment 2 is *Bright Beginnings*.

SOURCE: PCER Parent Interview (Fall 2003, Spring 2004, and Spring 2005).

The demographic characteristics of the primary caregivers, who were most often the biological or adoptive mother, are presented in table 2.2. The average age of the primary caregiver was 31 years. More than half (65%) of the primary caregivers were married. Less than half reported having had some college (34%) or had graduated from college (7%), 38 percent had a high school diploma or GED, and 21 percent had not finished high school. Less than half (43%) of the primary caregivers were employed full-time, 13 percent were employed part-time, and 39 percent were unemployed. There were no statistically detectable differences between the treatment and control groups on the primary caregiver characteristics.

Table 2.2. Primary caregiver demographic characteristics for *Bright Beginnings* and *Creative Curriculum*

Characteristics	Full sample n = 249	Curriculum comparison		
		Control n = 78	Treatment 1 ¹ n = 78	Treatment 2 ² n = 93
Age at baseline (years), mean	30.6	31.3	29.5	31.0
Marital status (%)				
Married	64.7	66.7	61.5	65.6
Separated/Divorced	18.5	17.9	17.9	19.4
Widowed	‡	‡	0.0	‡
Never Married	15.7	14.1	20.5	12.9
Race/ethnicity (%)				
White, non-Hispanic	83.5	93.6	76.9	80.4
African American, non-Hispanic	7.3	3.8	10.3	7.6
Hispanic	7.7	2.6	11.5	8.7
Asian or Pacific Islander	‡	0.0	0.0	‡
Native American	‡	0.0	0.0	‡
Multiple/other	‡	0.0	‡	‡
Educational level (%)				
Did not finish high school	21.4	12.8	25.6	25.0
High school diploma or GED	38.3	50.0	34.6	31.5
Some college	33.5	34.6	35.9	30.4
College graduate	6.9	‡	‡	13
Employment (%)				
Full-time	43.4	42.3	39.7	47.3
Part-time	12.9	16.7	10.3	11.8
Unemployed	39.0	37.2	44.9	35.5
Other	4.8	‡	5.1	5.4

‡ Reporting standards not met. Values suppressed to protect participant confidentiality.

¹Treatment 1 is *Creative Curriculum*.

²Treatment 2 is *Bright Beginnings*.

SOURCE: PCER Parent Interview (Fall 2003, Spring 2004, and Spring 2005).

Teachers

There were 21 teachers who participated in the preschool year intervention study. All of the teachers were female, and all were White. On average, the preschool teachers had 11 years of teaching experience, with an average of 6 years of experience teaching preschool. All of the teachers had a bachelor's (52%) or graduate (48%) degree. All reported having a state-awarded teacher certification. Table 2.3 provides additional information on the characteristics of the preschool sample of teachers. There were no statistically detectable differences between the treatment and control groups on the teacher characteristics.

Table 2.3. Preschool teacher characteristics for *Bright Beginnings* and *Creative Curriculum*

Characteristics	Full sample n = 21	Curriculum comparison		
		Control n = 7	Treatment 1 ¹ n = 7	Treatment 2 ² n = 7
Gender (% female)	100.0	100.0	100.0	100.0
Race/ethnicity (%)				
White, non-Hispanic	100.0	100.0	100.0	100.0
African American, non-Hispanic	0.0	0.0	0.0	0.0
Hispanic	0.0	0.0	0.0	0.0
Asian or Pacific Islander	0.0	0.0	0.0	0.0
Native American	0.0	0.0	0.0	0.0
Multiple/other	0.0	0.0	0.0	0.0
Educational level (%)				
High school diploma or GED	0.0	0.0	0.0	0.0
Associate's degree	0.0	0.0	0.0	0.0
Bachelor's degree	52.0	‡	‡	71.0
Graduate degree	48.0	57.0	57.0	‡
Current teaching license/certificate (%)	100.0	100.0	100.0	100.0
Child Development Associate (CDA) (%)	0.0	0.0	0.0	0.0
State-awarded preschool certificate (%)	0.0	0.0	0.0	0.0
No credential (%)	0.0	0.0	0.0	0.0
Years of teaching experience, overall (mean)	11.2	7.6	11.4	14.7
Years of preschool teaching experience (mean)	5.9	3.4	4.5	9.6

‡ Reporting standards not met. Values suppressed to protect participant confidentiality.

¹Treatment 1 is *Creative Curriculum*.

²Treatment 2 is *Bright Beginnings*.

SOURCE: PCER Preschool Teacher Survey (Fall 2003 and Spring 2004).

Programs/Classrooms

The average preschool class size was 16.6 children. The child-staff ratio was on average 7.2 children to 1 teacher or program staff person.

Random Assignment

Randomization of 36 classrooms to the three curriculum conditions was done for the site-specific evaluation during the pilot year of curriculum implementation (2002-03). During the pilot year, 21 of those classrooms were randomly selected to also participate in the PCER initiative. For the second year of implementation (2003-04), 13 of those 21 classrooms continued to participate in the initiative. These included five

implementing *Bright Beginnings*, four implementing *Creative Curriculum*, and four control classrooms. The 8 classrooms that dropped out were replaced by other classrooms (randomly selected) from the original 36 classrooms. These included two implementing *Bright Beginnings*, three implementing *Creative Curriculum*, and three control classrooms. All the teachers in these 21 classrooms were the same as during the pilot year, except for one teacher in a control classroom, and she had substituted for the teacher in that classroom while the latter was on maternity leave.

For the initial randomization, a total of 36 state pre-kindergarten classrooms within 28 schools in 7 county school systems were assigned to treatment and control conditions at the beginning of the pilot year of the study. The preschool classrooms were blocked into groups of three by matching them as closely as possible on demographic and academic performance variables for the elementary school geographically nearest each preschool (in many cases the preschool was in the same school facility as the elementary school). The matching variables were derived from information available on the Tennessee State Department of Education website and consisted of two composite factors that were created for this purpose—a demographic factor (urban/rural and percentage of races other than White) and a composite achievement factor (percent free lunch and reading, language, mathematics, and science achievement test scores). Classrooms in the same school—no more than two in any instance—were included as a single unit in these blocks to ensure that they would not be assigned to different conditions. Within each block, one classroom (or pair, if two in the same school were a single unit in the block) was randomly assigned to the *Bright Beginnings* curriculum condition, one to the *Creative Curriculum* condition, and one to the control group condition, with the constraint that the classrooms in a given county school system be distributed over the three conditions. All 36 teachers and their assistants consented to participate in the pilot-year study. The evaluation for the PCER initiative was conducted on a subset of the larger site-specific sample of teachers, parents, and children. That subset consisted of a random selection of 21 of the 36 classrooms and 309 children in those classrooms.

Contamination

The 21 pre-kindergarten classrooms in the evaluation were divided across 18 schools. For 15 of those schools, there was only one classroom that participated in the study. For three schools, two classrooms participated, with both assigned to the same experimental condition. Consequently, the risk of contamination from teachers in different conditions exchanging curriculum information or materials was minimal.

Control Condition

In the control condition, teachers used teacher-developed, nonspecific curricula with a focus on basic school readiness.

Data Collection

RTI International (RTI) collected the child, teacher, and school data for the Tennessee site for all three waves of data collection. The Tennessee research team was responsible for conducting the parent interviews in the preschool year. In the kindergarten follow-up year, RTI staff completed the parent interviews.

The fall assessment data collection window for child assessments ranged from September 4, 2003 to November 7, 2003. The average delay from the beginning of the treatment (i.e., start of the school year) to the beginning of the fall assessment window was 8 days. The spring pre-kindergarten window was March 30, 2004 to May 11, 2004, and the kindergarten follow-up window was April 4, 2005 to June 24, 2005.

Attrition

Twenty-one classrooms were randomly assigned to treatment or control condition, all of which remained in the study throughout the pre-kindergarten year. For the child assessment, the baseline (fall 2003) response rate was 100 percent, the spring 2004 response rate was 95 percent, and the kindergarten follow-up response rate was 98 percent.

Implementation

The teacher sample included teachers who participated in the pilot year of the study (2002-03), and new teachers who started in 2003-04. The intervention teachers received 2.5 days of curriculum implementation training prior to the beginning of the school year. Teachers had access to ongoing curriculum implementation support throughout the school year. Onsite consultation to teachers was provided four times during the school year—twice by trained Tennessee research staff members, and twice by curriculum trainers. Consultation visits typically included a classroom observation, an opportunity for teachers to ask questions about the curriculum, and implementation feedback from the curriculum trainer and/or research assistant. Consultation visits were conducted in September, October, November to January, and late February.

The research team conducted site-specific curriculum fidelity classroom observations three times throughout the year in the fall (October to November 2003), winter (January to February 2004), and spring (April to May 2004). Visits were made to both treatment and control classrooms. Each classroom was visited for a full morning, followed by an interview with the teacher. The developers of each curriculum provided the researchers with a fidelity instrument used to assess implementation. Both instruments were used in all the classrooms, including the control classrooms. Both of the site-specific fidelity measures included items that addressed general early childhood practice, as well as items that focused on specific activities/practices/materials that were unique to each curriculum. For the *Bright Beginnings* intervention curriculum, all classrooms were rated on a four-point scale (weak, fair, good, excellent). The site-specific fidelity measures indicated that most classrooms were rated as showing a High or Medium level of implementation. One classroom was rated as Low on the fidelity measure. The *Creative Curriculum* classrooms were rated as High (2), Medium (2), Low (2), and Not at All (1) on the site-specific fidelity measure. The control classrooms received an average rating of Medium on the fidelity measure.

Implementation Fidelity Ratings

Bright Beginnings

Each research team used a global fidelity measure to rate the overall fidelity with which the curricula were implemented in the preschool year of the project. A four-point scale ranging from “Not at All” (0) to “High” (3) was used to rate each treatment classroom. Researchers were asked to use their site-specific implementation and fidelity data to rate each treatment classroom on the global fidelity measure as High, Medium, Low, or Not at All. Researchers were also asked to provide a global rating for the control group curriculum. *Bright Beginnings* was rated Medium (1.88) on the global implementation fidelity measure. The control group curriculum was rated at the Medium (2.0) level as well.

Creative Curriculum

Each research team used a global fidelity measure to rate the overall fidelity with which the curricula were implemented in the preschool year of the project. A four-point scale ranging from “Not at All” (0) to “High” (3) was used to rate each treatment classroom. Researchers were asked to use their site-specific implementation and fidelity data to rate each treatment classroom on the global fidelity measure as High, Medium, Low, or Not at All. Researchers were also asked to provide a global rating for the control group curriculum. *Creative Curriculum* was rated Medium (2.14) on implementation fidelity. The control group curriculum was rated at the Medium (2.0) level as well.

Impact Analysis Results

We present analyses for each curriculum separately, beginning with the analyses of the child-level measures (i.e., the mathematics, reading, phonological awareness, and language assessments) followed by the analyses of the classroom observation data.

Bright Beginnings—Child Outcomes

The unadjusted mean scores for the child-level measures are reported in table C-1a in appendix C. Covariate adjusted mean differences and standard errors are presented in table D-1a in appendix D. For all analyses of the measures, the following covariates were included: (a) child's age, (b) gender, (c) race/ethnicity, (d) disability status as reported by parent, and (e) mother's education. The student-level effect sizes (ESs) are presented in table 2.4.

Mathematics assessments

We used repeated measures linear spline models to analyze the data from all three mathematics measures (Woodcock Johnson [WJ] Applied Problems, Child Math Assessment-Abbreviated [CMA-A] Composite Score, and Shape Composition). There were no statistically significant differences on these measures for the fall assessment.

There were no statistically detectable differences on the spring pre-kindergarten or kindergarten assessments on any of the mathematics assessments.

Based on the analyses for the three mathematics measures we conclude that *Bright Beginnings* did not have an effect on mathematics relative to the control condition.

Reading assessments

Data from the three reading measures (Test of Early Reading Ability [TERA], WJ Letter Word Identification, and WJ Spelling) were analyzed using repeated measures linear spline models. There were no statistically significant differences on these measures for the fall assessment.

There were no statistically detectable differences in the spring pre-kindergarten or kindergarten assessments on the WJ Letter Word Identification test or WJ Spelling test.

In spring of the pre-kindergarten year, there was a statistically reliable mean difference in scores on the TERA (ESs = .39, $p < .05$) favoring the *Bright Beginnings* group. No difference was found for the TERA for the spring kindergarten assessment.

Based on the analyses for the three reading measures, we conclude that *Bright Beginnings* did not have an effect on young children's early reading skills relative to the control condition.

Phonological awareness

The phonological awareness measures were the Preschool Comprehensive Test of Phonological and Print Processing (Pre-CTOPPP), Elision subtest, and the Comprehensive Test of Phonological Processing (CTOPP), Kindergarten, Elision subtest. We conducted a repeated measures analysis on the Pre-CTOPPP fall and spring pre-kindergarten data. There was no statistically significant difference on the Pre-CTOPPP for the fall assessment.

There was no statistically detectable difference on the Pre-CTOPPP for the spring pre-kindergarten assessment.

We analyzed the kindergarten CTOPP data using analysis of covariance (ANCOVA). For the ANCOVA analysis, the covariates were (a) the Pre-CTOPPP fall assessment score, (b) child's gender, (c) age, (d) race/ethnicity, (e) disability status as reported by parent, and (f) mother's education. There was no statistically significant difference between groups on the CTOPP for the spring kindergarten assessment.

Based on the analyses of the Pre-CTOPPP and CTOPP, we conclude that *Bright Beginnings* did not have an effect on phonological awareness relative to the control condition.

Language assessments

Data from the two language measures (Peabody Picture Vocabulary Test [PPVT] and Test of Language Development [TOLD] Grammatical Understanding subtest) were analyzed using repeated measures linear spline models. There were no statistically significant differences on either measure for the fall assessment.

There were no statistically detectable differences on any of these measures for the spring pre-kindergarten or spring kindergarten assessments.

Based on the analyses of the two language measures, we conclude that *Bright Beginnings* did not have an effect on language development relative to the control condition.

Behavioral outcomes

We conducted a repeated measures analysis for all three pre-kindergarten social behavioral measures (Social Skills Rating System [SSRS] Social Skills scale, SSRS Problem Behaviors scale, and Preschool Learning Behaviors Scale [PLBS]). The covariates were (a) child's age, (b) gender, (c) race/ethnicity, (d) disability status as reported by the parent, and (e) mother's education. There were no statistically significant differences on these measures for the fall assessment.

For the spring pre-kindergarten assessment, there were no statistically detectable differences on any of these measures.

We analyzed the data from the kindergarten versions of the three behavioral measures (SSRS Social Skills scale, SSRS Problem Behaviors scale, and Learning Behaviors Scale [LBS]) using analysis of covariance (ANCOVA). For the ANCOVA analyses, the covariates included the fall pre-kindergarten score of the pre-kindergarten version of the relevant test, along with (a) child's age, (b) gender, (c) race/ethnicity, (d) disability status as reported by the parent, and (e) mother's education.

There were no statistically detectable differences between groups on any of these measures for the spring kindergarten assessment.

Based on the analyses of the three behavioral measures, we conclude that *Bright Beginnings* did not have an effect on children's social and learning behaviors relative to the control condition.

Bright Beginnings—Classroom Outcomes

The unadjusted mean scores for classroom measures are reported in table C-1b in appendix C. Covariate adjusted mean differences and standard errors are presented in table D-1b in appendix D. For all analyses of classroom measures, the following variables were included in the model as covariates: (a) teacher has a BA degree, (b) previous teaching experience, (c) child/adult ratio in classroom, (d) average class size, (e) city size, and (f) geographic site. The classroom-level effect sizes (ES_c) are presented in table 2.4.

Overall classroom environment

We conducted a repeated measures analysis on the Early Childhood Environment Rating Scale-Revised (ECERS-R). There was a statistically significant difference between groups for the fall observation ($ES_c = 1.39$, $p < .05$; follow-up analyses for this finding are included in appendix A). *Bright Beginnings* classrooms received higher global classroom quality ratings.

No statistically detectable differences between groups were obtained for the spring pre-kindergarten observation.

Based on the analysis of the ECERS-R, we conclude that *Bright Beginnings* did not have an effect on overall classroom quality relative to the control condition.

Teacher-child relationships

We obtained observations on the Arnett Detachment, Harshness, Permissiveness, and Positive Interactions scales in fall and spring of the pre-kindergarten year, and conducted repeated measures analyses. There was a

statistically significant difference at the time of the fall observation on the Arnett Detachment scale ($ES_c = -1.16, p < .05$; follow-up analyses for this finding are included in appendix A). Teachers in the *Bright Beginnings* classrooms were rated as being less detached in their interactions with their students relative to teachers in the control classrooms. There were no statistically significant differences on the other scales for the fall pre-kindergarten observation.

In spring of the kindergarten year, there were no statistically detectable differences between groups on the Arnett Detachment, Harshness, Permissiveness, or Positive Interactions scales.

Based on the analyses of the four Arnett scales, we conclude that *Bright Beginnings* did not have an effect on teacher-child relationships relative to the control condition.

Classroom instruction

We obtained observations on classroom instruction in (a) early literacy (Teacher Behavior Rating Scale [TBRS] Print and Letter Knowledge and TBRS Written Expression scales); (b) phonological awareness (TBRS Phonological Awareness scale); (c) language (TBRS Book Reading and Oral Language scales); and (d) early mathematics (TBRS Math Concepts scale) in spring of the pre-kindergarten year only. To analyze these data, ANCOVAs were conducted; the covariates were: (a) teacher has a BA degree, (b) previous teaching experience, (c) child/adult ratio in classroom, (d) average class size, (e) city size, and (f) geographic site.

There were statistically reliable effects favoring the *Bright Beginnings* classrooms on the Phonological Awareness ($ES_c = 1.53, p < .05$), Print and Letter Knowledge ($ES_c = 1.51, p < .05$), and Written Expression ($ES_c = 1.61, p < .01$) scales of the TBRS. There were no statistically detectable differences on the remaining scales.

Based on the analyses of the TBRS Print and Letter Knowledge and Written Expression scales, we conclude that *Bright Beginnings* had a positive effect on early literacy and phonological awareness instruction relative to the control condition.

Based on the analysis of the TBRS Phonological Awareness scale, we conclude that *Bright Beginnings* had a positive effect on instruction in phonological awareness relative to the control condition.

Based on the analysis of the TBRS Book Reading and Oral Language scales, we conclude that *Bright Beginnings* did not have a statistically detectable effect on language instruction.

Based on the analysis of the TBRS Math Concepts scale, we conclude that *Bright Beginnings* did not have a statistically detectable effect on early mathematics instruction.

Summary of Findings for *Bright Beginnings*

The findings for *Bright Beginnings* are summarized in table 2.4.

Table 2.4. Effect sizes for *Bright Beginnings*

Measure	Student-level effect sizes (ES _s)		
	RM analysis	RM analysis	ANCOVA
	Spring Pre-K	Spring K	Spring K
Mathematics			
WJ Applied Problems	.16	.13	—
CMA-A Mathematics Composite	.14	.07	—
Shape Composition ¹	-.03	.15	—
Reading			
TERA	.39*	-.07	—
WJ Letter Word Identification	.35	.09	—
WJ Spelling	.18	.06	—
Phonological awareness			
Pre-CTOPPP/CTOPP	-.07	†	.01
Language			
PPVT	.13	.07	—
TOLD	.09	.16	—
Behavior			
SSRS Social Skills	-.27	†	-.03
SSRS Problem Behavior ²	.23	†	.24
PLBS/LBS	.04	†	-.30
	Classroom-level effect sizes (ES _c)		
Measure	RM analysis	ANCOVA	
	Spring Pre-K	Spring Pre-K	
Global classroom quality			
ECERS-R	.80	—	
Teacher-child interaction			
Arnett Detachment ³	.19	—	
Arnett Harshness ³	.12	—	
Arnett Permissiveness ³	.16	—	
Arnett Positive Interactions	.41	—	
Teacher instructional practices ⁴			
TBRS Book Reading	†	1.03	
TBRS Oral Language	†	.39	
TBRS Phonological Awareness	†	1.53*	
TBRS Print and Letter Knowledge	†	1.51*	
TBRS Written Expression	†	1.61*	
TBRS Math Concepts	†	.98	

— Not available.

† Not applicable. Four of the kindergarten student-level measures were not on the same scale as the pre-kindergarten measures. The classroom-level data were only collected during the pre-kindergarten year of the study.

* $p < .05$ ¹ Building Blocks, Shape Composition task² Higher scores on this scale represent more negative child behaviors.³ Lower scores on this scale represent a more positive classroom environment.⁴ ANCOVA models for the TBRS measures did not include baseline pretest scores because TBRS data were only collected in spring of the pre-kindergarten year.

NOTE: RM: Repeated Measures

ANCOVA: Analysis of covariance

Significance indications (p -values) in the table refer to the tests of contrasts between intervention and control groups that underlie the effect sizes reported here. Refer to the glossary for abbreviations of the measures.

SOURCE: The Preschool Curriculum Evaluation Research (PCER) Study.

Creative Curriculum—Child Outcomes

The unadjusted mean scores for child-level measures are reported in table C-2a in appendix C. Covariate adjusted mean differences and standard errors are reported in table D-2a in appendix D. For all analyses of child-level measures, the following covariates were included: (a) child's age, (b) gender, (c) race/ethnicity, (d) disability status as reported by parent, and (e) mother's education. The student-level effect sizes (ESs) are presented in table 2.5.

Mathematics assessments

We used repeated measures linear spline models to analyze the data from all three mathematics measures (WJ Applied Problems, CMA-A Composite Score, and Shape Composition). There were no statistically detectable differences at the fall assessment.

There were no statistically detectable differences on the spring pre-kindergarten or kindergarten assessments on any of the mathematics assessments.

Based on the analyses for the three mathematics measures, we conclude that *Creative Curriculum* did not have a statistically detectable effect on mathematics relative to the control condition.

Reading assessments

Data from the three reading measures (TERA, WJ Letter Word Identification, and WJ Spelling) were analyzed using repeated measures linear spline models. There were no statistically detectable differences at the fall assessment.

There were no statistically detectable differences on any of these measures for the spring pre-kindergarten or spring kindergarten assessments.

Based on the analyses for the three reading measures, we conclude that *Creative Curriculum* did not have a statistically detectable effect on early reading relative to the control condition.

Phonological awareness

The phonological awareness measures were the Pre-CTOPPP, Elision subtest, and the CTOPP, Kindergarten, Elision subtest. We conducted a repeated measures analysis on the Pre-CTOPPP fall and spring pre-kindergarten data. There was no statistically detectable difference on the Pre-CTOPPP for the fall assessment.

There was no statistically detectable difference on the Pre-CTOPPP for the spring pre-kindergarten assessment.

We analyzed the kindergarten CTOPP data using analysis of covariance (ANCOVA). For the ANCOVA analysis, the covariates were the (a) Pre-CTOPPP fall assessment score, (b) child's gender, (c) age, (d) race/ethnicity, (e) disability status as reported by parent, and (f) mother's education. There was no statistically detectable difference between groups on the CTOPP for the spring kindergarten assessment.

Based on the analyses of the Pre-CTOPPP and CTOPP, we conclude that *Creative Curriculum* did not have a statistically detectable effect on phonological awareness relative to the control condition.

Language assessments

Data from the two language measures (Peabody Picture Vocabulary Test [PPVT] and Test of Language Development [TOLD] Grammatical Understanding subtest) were analyzed using repeated measures linear spline models. There were no statistically detectable differences for the fall assessment.

There were no statistically detectable differences on either of these measures for the spring pre-kindergarten or spring kindergarten assessments.

Based on the analyses for the two language measures, we conclude that *Creative Curriculum* did not have an effect on language development relative to the control condition.

Behavioral outcomes

We conducted a repeated measures analysis for all three pre-kindergarten social behavioral measures (SSRS Social Skills scale, SSRS Problem Behaviors scale, and Preschool Learning Behaviors Scale [PLBS]). The covariates were (a) child's age, (b) gender, (c) race/ethnicity, (d) disability status as reported by the parent, and (e) mother's education. There were no statistically detectable differences on these measures for the fall assessment.

For the spring pre-kindergarten assessment, there were no statistically detectable differences on any of these measures.

We analyzed the data from the kindergarten versions of the three behavioral measures (SSRS Social Skills scale, SSRS Problem Behaviors scale, Learning Behaviors Scale [LBS]) using analysis of covariance (ANCOVA). For the ANCOVA analyses, the covariates included (a) the fall pre-kindergarten score of the pre-kindergarten version of the relevant test, along with (b) child's age, (c) gender, (d) race/ethnicity, (e) disability status as reported by the parent, and (f) mother's education.

There were no statistically detectable differences between groups on any of these measures for the spring kindergarten assessment.

Based on the analyses of the three behavioral measures, we conclude that *Creative Curriculum* did not have an effect on children's social and learning behaviors relative to the control condition.

Creative Curriculum—Classroom Outcomes

The unadjusted mean scores for the classroom-level measures are reported in table C-2b in appendix C. Covariate adjusted mean differences and standard errors are reported in table D-2b in appendix D. For all analyses of the classroom measures, the following variables were included in the model as covariates: (a) teacher has a BA degree, (b) previous teaching experience, (c) child/adult ratio in classroom, (d) average class size, (e) city size, and (f) geographic site. The classroom-level effect sizes (ES_C) are presented in table 2.5.

Overall classroom environment

We conducted a repeated measures analysis on the ECERS-R. There was a statistically significant difference between groups on the fall observation ($ES_C = 1.94, p < .01$); follow-up analyses for this finding are included in appendix A. *Creative Curriculum* teachers received higher overall classroom quality ratings.

No statistically detectable difference between groups was obtained for the spring pre-kindergarten observation.

Based on the analysis of the ECERS-R, we conclude that *Creative Curriculum* did not have a statistically detectable effect on overall classroom quality relative to the control condition.

Teacher-child relationships

We obtained observations on the Arnett Detachment, Harshness, Permissiveness, and Positive Interactions scales in fall and spring of the pre-kindergarten year, and conducted repeated measures analyses. There was a statistically significant difference at the time of the fall observation on the Arnett Detachment scale ($ES_C = -.95, p < .05$; follow-up analyses for this finding are included in appendix A). Teachers in the *Creative Curriculum* classrooms were rated as being less detached in their interactions with their students relative to teachers in the control classrooms. There were no statistically detectable differences on the other scales for the fall observation

In spring of the pre-kindergarten year, there were no statistically significant differences between groups on the Arnett Detachment, Harshness, Permissiveness, or Positive Interaction scales.

Based on the analyses of the four Arnett scales, we conclude that *Creative Curriculum* did not have a statistically detectable effect on teacher-child relationships relative to the control condition.

Classroom instruction

We obtained observations on classroom instruction in (a) early literacy (TBRS Print and Letter Knowledge and Written Expression scales); (b) phonological awareness (TBRS Phonological Awareness scale); (c) language (TBRS Book Reading and Oral Language scales); and (d) early mathematics (TBRS Math Concepts scale) in spring of pre-kindergarten only. To analyze these data, ANCOVAs were conducted; the covariates were: (a) teacher has a BA degree, (b) previous teaching experience, (c) child/adult ratio in classroom, (d) average class size, (e) city size, and (f) geographic site.

There were no statistically detectable differences between the treatment and control classrooms on the TBRS scales.

Based on the analyses of the TBRS Print and Letter Knowledge and Written Expression scales, we conclude that *Creative Curriculum* did not have a statistically detectable effect on early literacy instruction relative to the control condition.

Based on the analysis of the TBRS Phonological Awareness scale, we conclude that *Creative Curriculum* did not have a statistically detectable effect on instruction in phonological awareness relative to the control condition.

Based on the analysis of the TBRS Book Reading and Oral Language scales, we conclude that *Creative Curriculum* did not have a statistically detectable effect on language instruction.

Based on the analysis of the TBRS Math Concepts scale, we conclude that *Creative Curriculum* did not have a statistically detectable effect on early mathematics instruction.

Summary of Findings for *Creative Curriculum*

The findings for *Creative Curriculum* are summarized in table 2.5.

Table 2.5. Effect sizes for *Creative Curriculum: Tennessee*

Measure	Student-level effect sizes (ES _s)		
	RM analysis	RM analysis	ANCOVA
	Spring Pre-K	Spring K	Spring K
Mathematics			
WJ Applied Problems	.17	.17	—
CMA-A Mathematics Composite	.10	.05	—
Shape Composition ¹	-.12	.00	—
Reading			
TERA	.02	.10	—
WJ Letter Word Identification	.16	.38	—
WJ Spelling	.19	.25	—
Phonological awareness			
Pre-CTOPPP/CTOPP	-.10	†	.06
Language			
PPVT	.23	.12	—
TOLD	.07	.11	—
Behavior			
SSRS Social Skills	-.03	†	.35
SSRS Problem Behavior ²	.07	†	-.05
PLBS/LBS	.14	†	.08
Measure	Classroom-level effect sizes (ES _c)		
	RM analysis	ANCOVA	
	Spring Pre-K	Spring Pre-K	
Global classroom quality			
ECERS-R	.45	—	
Teacher-child interaction			
Arnett Detachment ³	-.16	—	
Arnett Harshness ³	-.12	—	
Arnett Permissiveness ³	.51	—	
Arnett Positive Interactions	-.15	—	
Teacher instructional practices ⁴			
TBRS Book Reading	†	-.47	
TBRS Oral Language	†	-.07	
TBRS Phonological Awareness	†	1.97	
TBRS Print and Letter Knowledge	†	1.81	
TBRS Written Expression	†	1.99	
TBRS Math Concepts	†	1.48	

— Not available.

† Not applicable. Four of the kindergarten student-level measures were not on the same scale as the pre-kindergarten measures. The classroom-level data were only collected during the pre-kindergarten year of the study.

¹ Building Blocks, Shape Composition task² Higher scores on this scale represent more negative child behaviors.³ Lower scores on this scale represent a more positive classroom environment.⁴ ANCOVA models for the TBRS measures did not include baseline pretest scores because TBRS data were only collected in spring of the pre-kindergarten year.

NOTE: RM: Repeated Measures

ANCOVA: Analysis of covariance

Significance indications (*p*-values) in the table refer to the tests of contrasts between intervention and control groups that underlie the effect sizes reported here. Refer to the glossary for abbreviations of the measures.

SOURCE: The Preschool Curriculum Evaluation Research (PCER) Study.

Chapter 3. *Creative Curriculum*: University of North Carolina at Charlotte (North Carolina and Georgia sites)

Curriculum

The research team at the University of North Carolina at Charlotte (North Carolina), implemented *Creative Curriculum*, a comprehensive curriculum for 3- to 5-year-old children.² The curriculum addresses four areas of development: social/emotional, physical, cognitive, and language development.

Creative Curriculum requires the physical space of the classroom to be structured into 10 interest areas: blocks, dramatic play, toys and games, art, library, discovery, sand and water, music and movement, cooking, and computers. Time is also allotted for outdoor activities. The 10 interest areas are designed to address curriculum content, such as literacy, mathematics, science, social studies, the arts, technology, and process skills, such as observing, exploring, and problem solving.

Sample

The North Carolina research team recruited Head Start programs in North Carolina and Georgia. All of the programs were full-day programs. Head Start teachers, assistants, and site managers were offered a stipend for participating in the study. Eight classrooms in North Carolina and 10 classrooms in Georgia were recruited to participate in the study. The parental consent process began before the start of the school year. The North Carolina research team relied on teachers to assist them in recruiting parents for the study. Teachers were given a letter to give to the parents during their initial home visit with the parents. Any parents who did not participate in a home visit were given a letter when he or she first came to the school. A sample of 18 classrooms and 194 children (97 treatment, 97 control) and parents were recruited for participation in the study. Data were collected on 190 children and 168 parents at the time of the fall baseline data collection.

In the follow-up year of the evaluation, the sample of schools went from five in pre-kindergarten to more than 54 schools in kindergarten. The sample of classrooms went from 18 preschool to 122 kindergarten classrooms. The kindergarten sample included 190 children and parents from the original sample of participants. Data were collected on 162 children and 135 parents.

Children and Families

The children were 4.5 years old at the time of baseline data collection and less than half were male (46%). The majority of the children were African American (85%). Table 3.1 provides additional information on the demographic characteristics of the children in the study sample. There were no statistically detectable differences between the treatment and control groups on these child characteristics.

The demographic characteristics of the primary caregivers, who were most often the biological or adoptive mother, are presented in table 3.2. The average age of the primary caregiver was 32 years. Less than half (39%) were married, and 39 percent were never married. Almost half of the primary caregivers reported having had some college (42%) or had graduated from college (6%), 29 percent had a high school diploma or GED, and 23 percent had not finished high school. Less than half (45%) of the primary caregivers were employed full-time, 10 percent were employed part-time, and 41 percent were unemployed. There were no statistically detectable differences between the treatment and control groups on the primary caregiver characteristics.

² The Vanderbilt University (Tennessee) research team also evaluated *Creative Curriculum*.

Table 3.1. Child demographic characteristics for *Creative Curriculum*: North Carolina and Georgia

Characteristics	Full sample n = 194	Curriculum comparison	
		Control n = 97	Treatment n = 97
<i>Creative Curriculum: North Carolina and Georgia</i>			
Age at baseline (years), mean	4.5	4.5	4.5
Gender (% male)	45.8	47.4	44.2
Race/ethnicity (%)			
White, non-Hispanic	2.9	‡	‡
African American, non-Hispanic	85.0	84.1	85.9
Hispanic	7.5	6.8	8.2
Asian or Pacific Islander	‡	0.0	
Native American	0.0	0.0	0.0
Multiple/other	4.0	5.7	‡
Child disability status (parent reported, %)	20.1	23.9	16.0
Characteristics	Full sample n = 97	Curriculum comparison	
		Control n = 48	Treatment n = 49
<i>Creative Curriculum: North Carolina</i>			
Age at baseline (years), mean	4.5	4.5	4.5
Gender (% male)	46.4	47.9	44.9
Race/ethnicity (%)			
White, non-Hispanic	5.5	‡	‡
African American, non-Hispanic	81.3	80.9	81.8
Hispanic	11.0	10.6	11.4
Asian or Pacific Islander	0.0	0.0	0.0
Native American	0.0	0.0	0.0
Multiple/other	‡	‡	0.0
Child disability status (parent reported, %)	28.7	30.4	26.8
Characteristics	Full sample n = 93	Curriculum comparison	
		Control n = 47	Treatment n = 46
<i>Creative Curriculum: Georgia</i>			
Age at baseline (years), mean	4.5	4.5	4.5
Gender (% male)	45.2	46.8	43.5
Race/ethnicity (%)			
White, non-Hispanic	0.0	0.0	0.0
African American, non-Hispanic	89.0	87.8	90.2
Hispanic	‡	‡	‡
Asian or Pacific Islander	‡	0.0	‡
Native American	0.0	0.0	0.0
Multiple/other	6.1	9.8	‡
Child disability status (parent reported, %)	11.0	16.7	5.0

‡ Reporting standards not met. Values suppressed to protect participant confidentiality.

SOURCE: PCER Parent Interview (Fall 2003, Spring 2004, and Spring 2005).

Table 3.2. Primary caregiver demographic characteristics for *Creative Curriculum*: North Carolina and Georgia

Characteristics	Full sample n = 169	Curriculum comparison	
		Control n = 88	Treatment n = 81
Age at baseline (years), mean	31.8	32.4	31.1
Marital status (%)			
Married	38.5	42.0	34.6
Separated/Divorced	21.9	19.3	24.7
Widowed	‡	‡	0.0
Never Married	39.1	37.5	40.7
Race/ethnicity (%)			
White, non-Hispanic	4.1	‡	4.9
African American, non-Hispanic	81.7	83.0	80.2
Hispanic	7.1	5.7	8.6
Asian or Pacific Islander	‡	0.0	‡
Native American	0.0	0.0	0.0
Multiple/other	6.5	8.0	4.9
Educational level (%)			
Did not finish high school	22.8	20.9	24.7
High school diploma or GED	29.3	30.2	28.4
Some college	41.9	39.5	44.4
College graduate	6.0	9.3	‡
Employment (%)			
Full-time	45.0	42	48.1
Part-time	10.1	10.2	9.9
Unemployed	40.8	40.9	40.7
Other	4.1	6.8	‡

‡ Reporting standards not met. Values suppressed to protect participant confidentiality.

SOURCE: PCER Parent Interview (Fall 2003, Spring 2004, and Spring 2005).

Teachers

There were 18 teachers who participated in the preschool year intervention study. All of the teachers were female. The majority of the teachers were African American (89%). The preschool teachers had on average 12 years of teaching experience, with an average of 9 years of experience teaching preschool. Half (50%) had an associate's degree, and 44 percent had a high school diploma or GED. Seventy-two percent of the teachers had a state-awarded teacher certification, and 78 percent had a Child Development Associate (CDA) credential. Table 3.3 provides additional information on the characteristics of the preschool sample of teachers. There were no statistically detectable differences between the treatment and control groups on the teacher characteristics.

Programs/Classrooms

The average preschool class size was 14.4 children. The child-staff ratio was on average 7.6 children to one teacher or program staff person.

Table 3.3. Preschool teacher characteristics for *Creative Curriculum*: North Carolina and Georgia

Characteristics	Full sample n = 18	Curriculum comparison	
		Control n = 9	Treatment n = 9
Gender (% female)	100.0	100.0	100.0
Race/ethnicity (%)			
White, non-Hispanic	‡	‡	0.0
African American, non-Hispanic	89.0	78.0	100.0
Hispanic	0.0	0.0	0.0
Asian or Pacific Islander	0.0	0.0	0.0
Native American	0.0	0.0	0.0
Multiple/other	6.0	11.0	0.0
Educational level (%)			
High school diploma or GED	44.0	67.0	‡
Associate's degree	50.0	‡	67.0
Bachelor's degree	‡	0.0	11.0
Graduate degree	0.0	0.0	0.0
Current teaching license/certificate (%)	24.0	0.0	44.0
Child Development Associate (CDA) (%)	78.0	78.0	78.0
State-awarded preschool certificate (%)	72.0	67.0	8.0
No credential (%)	0.0	0.0	0.0
Years of teaching experience, overall (mean)	11.9	12.9	11.0
Years of preschool teaching experience (mean)	8.7	9.2	8.2

‡ Reporting standards not met. Values suppressed to protect participant confidentiality.

SOURCE: PCER Preschool Teacher Survey (Fall 2003 and Spring 2004).

Random Assignment

Randomization of teachers within centers was done during the pilot year (2002-03) of the study, and the same assignments were maintained for the second year (2003-04) of curriculum implementation. At the end of the pilot year, the North Carolina site retained eight (four treatment and four control) of the 10 classrooms. Two classrooms were dropped because they were funded by the state's More at Four program, had degreed teachers, and had problems with high rates of teacher attrition. The Georgia site retained 10 out of 10 classrooms.

Treatment and control classrooms were housed in the same centers. Teachers were randomly assigned to either the treatment or control condition. The teachers within each center were assigned to blocks based on educational level and teacher certification status. They were then randomly assigned to treatment and control conditions within blocks. Children were randomly assigned to classrooms within each center. The children were randomly assigned to classrooms by blocking within each center by gender, special needs status, and ethnicity. Children were then randomly assigned to classrooms within blocks.

Each of these randomization procedures was conducted by using a pseudorandom number-generating software program to assign a random number to each participant. A random seed was set each time a new batch of numbers was generated. The participants were then sorted by their random number to make assignments to conditions. This process ensured that each participant (teacher or child) within each block had the same probability of assignment to treatment or control conditions.

A total of five Head Start preschools in two states (North Carolina and Georgia) were purposefully selected in the fall of the preschool year by the research team. There were three participating centers in North Carolina and two in Georgia. Randomization to either *Creative Curriculum* or the control curriculum was carried out at the classroom level and at child levels as described above. A total of 18 classrooms and 194 children took part in the study.

Contamination

Treatment and control classrooms were housed in the same centers. Teachers within the Head Start centers worked closely together. There may have been a few instances where a treatment group teacher inadvertently shared aspects from the treatment curriculum content or training with a control teacher. The research team conducted focus groups of both groups of teachers and there were not many comments about sharing.

The site-specific implementation fidelity data suggests that the control group teachers were doing some of the activities on the *Creative Curriculum* fidelity checklist. However, many of the items focused on generally accepted early childhood practice and were not curriculum-specific. It is not possible to attribute the control group scores to a contamination effect.

Control Condition

In the control condition, teachers used teacher-developed, nonspecific curricula.

Data Collection

RTI International (RTI) collected the child, teacher, and school data for the North Carolina and Georgia sites for all three waves of data collection. The North Carolina research team was responsible for conducting the parent interviews in the preschool year. In the kindergarten follow-up year, RTI staff completed the parent interviews. The fall assessment data collection window for child assessments ranged from September 3, 2003 to October 24, 2003 in North Carolina and August 25, 2003 to October 15, 2003 in Georgia. The average delay from the beginning of the treatment (i.e., start of the school year) to the beginning of the fall assessment window was 16 days in North Carolina and 14 days in Georgia. The spring pre-kindergarten window was March 3, 2004 to May 11, 2004 in North Carolina and April 13, 2004 to June 19, 2004 in Georgia. The kindergarten follow-up window was April 4, 2005 to May 24, 2005 in North Carolina and March 28, 2005 to June 30, 2005 in Georgia.

Attrition

Eighteen classrooms were randomly assigned to treatment and control conditions. All 18 classrooms remained in the study throughout the pre-kindergarten year.

For the child assessment, the baseline (fall 2003) response rate was 98 percent, the spring 2004 response rate was 90 percent, and the kindergarten follow-up response rate was 85 percent.

Implementation

Teachers in the treatment group were in their second year of implementing *Creative Curriculum* at the time of the evaluation. The North Carolina research team provided refresher training to the treatment group teachers during the second year of curriculum implementation. The training was delivered in 1-day or half-day sessions that were offered between August 2003 and February 2004. Four training sessions were provided to the treatment group teachers in North Carolina. Five training sessions were provided to the treatment group teachers in Georgia. Training topics included choosing and planning in-depth topics of study with children; providing materials and interactions for content learning (literacy, mathematics, science, social studies, the arts, and technology); and observation-based assessment of children's learning. The training was designed as a mix of lecture, small group projects, video viewing, and hands-on practical application activities. The same

trainer conducted training in both North Carolina and Georgia. Technical assistance was provided to teachers on an ongoing basis throughout the school year. The *Creative Curriculum* trainer and the local site coordinators at each project location provided technical assistance. Technical assistance was provided from August/September 2003 to April 2004.

An independent observer collected implementation fidelity data in fall of the preschool year. Observations were conducted in both treatment and control classrooms. The fidelity measure domains included the physical environment of the classrooms, the structure of the classrooms, teacher-child interactions, assessment, and family involvement. On average, the treatment group scored 86 percent on the implementation checklist and the control group scored approximately 58 percent on the implementation checklist. A *Creative Curriculum* classroom is considered as meeting the publishers' implementation criteria if the scores are greater than or equal to 80 percent at the scale level (e.g., teacher-child interactions or assessment) and 85 percent on the total score. At the time of the fall 2003 classroom fidelity observations, only two of nine treatment classrooms reached this level of implementation. None of the control classrooms met the criteria. The spring 2004 fidelity observations indicated that six of nine treatment classrooms had reached the implementation criteria. None of the control classrooms reached the 85 percent total score criterion.

Implementation Fidelity Ratings

Each research team used a global fidelity measure to rate the overall fidelity with which the curricula were implemented in the preschool year of the project. A four-point scale ranging from "Not at All" (0) to "High" (3) was used to rate each treatment classroom. Researchers were asked to use their site-specific implementation and fidelity data to rate each treatment classroom on the global fidelity measure as High, Medium, Low, or Not at All. Researchers were also asked to provide a global rating for the control group curriculum. *Creative Curriculum* was rated Medium (2.11) on the global implementation fidelity measure. The control group curriculum was rated at the low Medium (1.5) level.

Impact Analysis Results

We present analyses of the child-level measures (i.e., the mathematics, reading, phonological awareness, language, and behavioral assessments) followed by the analyses of the classroom observation data. Our discussion of the results focuses on the combined analyses of the Georgia and North Carolina sites.

Creative Curriculum—Child Outcomes

The unadjusted mean scores for the child-level measures are reported in table C-3a in appendix C. Covariate adjusted mean differences and standard errors are reported in table D-3a in appendix D. For all analyses of the child-level measures, the following covariates were included: (a) child's age, (b) gender, (c) race/ethnicity, (d) disability status as reported by parent, and (e) mother's education. The student-level effect sizes (ESs) are presented in table 3.4.

Mathematics assessments

We used repeated measures linear spline models to analyze the data from all three mathematics measures (Woodcock Johnson [WJ] Applied Problems, Child Math Assessment-Abbreviated [CMA-A] Composite Score, and Shape Composition). There were no statistically detectable differences for the fall assessment.

There were no statistically detectable differences on the spring pre-kindergarten or kindergarten assessments on any of the mathematics assessments.

Based on the analyses of the three mathematics measures, we conclude that *Creative Curriculum* did not have a statistically detectable effect on mathematics relative to the control condition.

Reading assessments

Data from the three reading measures (Test of Early Reading Ability [TERA], WJ Letter Word Identification, and WJ Spelling) were analyzed using repeated measures linear spline models. There were no statistically detectable differences for the fall assessment.

There were no statistically detectable differences on any of these measures for the spring pre-kindergarten or spring kindergarten assessments.

Based on the analyses of the three reading measures, we conclude that *Creative Curriculum* did not have a statistically detectable effect on reading relative to the control condition.

Phonological awareness

The phonological awareness measures were the Preschool Comprehensive Test of Phonological and Print Processing (Pre-CTOPPP), Elision subtest, and the Comprehensive Test of Phonological Processing (CTOPP), Kindergarten, Elision subtest. We conducted a repeated measures analysis on the Pre-CTOPPP fall and spring pre-kindergarten data. There was no statistically detectable difference on the Pre-CTOPPP for the fall assessment.

There was no statistically detectable difference on the Pre-CTOPPP for the spring pre-kindergarten assessment.

We analyzed the kindergarten CTOPP data using analysis of covariance (ANCOVA). For the ANCOVA analysis, the covariates were (a) the Pre-CTOPPP fall assessment score, (b) child's gender, (c) age, (d) race/ethnicity, (e) disability status as reported by parent, and (f) mother's education. There was no statistically detectable difference between groups on the CTOPP for the spring kindergarten assessment.

Based on the analyses of the Pre-CTOPPP and CTOPP, we conclude that *Creative Curriculum* did not have a statistically detectable effect on phonological awareness relative to the control condition.

Language assessments

Data from the two language measures (Peabody Picture Vocabulary Test [PPVT] and Test of Language Development [TOLD] Grammatical Understanding subtest) were analyzed using repeated measures linear spline models. There were no statistically detectable differences for the fall assessment.

There were no statistically detectable differences on any of these measures for the spring pre-kindergarten or spring kindergarten assessments.

Based on the analyses of the two language measures, we conclude that *Creative Curriculum* did not have a statistically detectable effect on language development relative to the control condition.

Behavioral outcomes

We conducted a repeated measures analysis for all three pre-kindergarten social behavioral measures (Social Skills Rating System [SSRS] Social Skills scale, SSRS Problem Behaviors scale, and Preschool Learning Behaviors Scale [PLBS]). The covariates were (a) child's age, (b) gender, (c) race/ethnicity, (d) disability status as reported by the parent, and (e) mother's education. There were no statistically detectable differences on these measures for the fall assessment.

For the spring pre-kindergarten assessment, there were no statistically detectable differences on any of these measures.

We analyzed the data from the kindergarten versions of the three behavioral measures (SSRS Social Skills scale, SSRS Problem Behaviors scale, and Learning Behaviors Scale [LBS]) using analysis of covariance (ANCOVA). For the ANCOVA analyses, the covariates included (a) the fall pre-kindergarten score of the pre-kindergarten version of the relevant test, along with (b) child's age, (c) gender, (d) race/ethnicity, (e) disability status as reported by the parent, and (f) mother's education.

There were no statistically detectable differences between groups on any of these measures for the spring kindergarten assessment.

Based on the analyses of the three behavioral measures, we conclude that *Creative Curriculum* did not have a statistically detectable effect on children's social and learning behaviors relative to the control condition.

Creative Curriculum—Classroom Outcomes

The unadjusted mean scores for classroom-level measures are reported in table C-3b in appendix C. Covariate adjusted mean differences and standard errors are reported in table D-3b in appendix D. For all analyses of classroom-level measures, the following variables were included in the model as covariates: (a) teacher has a BA degree, (b) previous teaching experience, (c) child/adult ratio in classroom, (d) average class size, (e) city size, and (f) geographic site. The classroom-level effect sizes (ES_C) are presented table 3.4.

Overall classroom environment

We conducted a repeated measures analysis on the Early Childhood Environment Rating Scale-Revised (ECERS-R). There was no statistically detectable difference between groups for the fall observation.

There was a statistically significant difference between the *Creative Curriculum* classrooms and the control classrooms on the ECERS-R for the spring pre-kindergarten observation ($ES_C = 1.66, p < .05$). Treatment group classrooms received higher overall classroom quality ratings relative to the control group classrooms.

Based on the analysis of the ECERS-R, we conclude that *Creative Curriculum* had a positive effect on overall classroom quality relative to the control condition.

Teacher-child relationships

We obtained observations on the Arnett Detachment, Harshness, Permissiveness, and Positive Interactions scales in fall and spring of the pre-kindergarten year, and conducted repeated measures analyses. There were no statistically detectable differences on these measures for the fall observation.

In spring of the pre-kindergarten year, there were statistically significant differences on the Arnett Detachment ($ES_C = -1.68, p < .05$) scale, indicating that teachers in the *Creative Curriculum* classrooms were less detached in their interactions with their students relative to teachers in the control classrooms. There was also a statistically significant effect on the Positive Interactions scale ($ES_C = 1.65, p < .01$), indicating that teachers in the *Creative Curriculum* classrooms were observed having more positive interactions with children relative to teachers in the control classrooms. No statistically detectable differences were obtained on the Arnett Harshness scale.

Based on the analyses of the four Arnett scales, we conclude that *Creative Curriculum* had a positive effect on teacher-child relationships relative to the control condition.

Classroom instruction

We obtained observations on classroom instruction in (a) early literacy (Teacher Behavior Rating Scale [TBRS] Print and Letter Knowledge and Written Expression scales), (b) phonological awareness (TBRS Phonological Awareness scale), (c) language (TBRS Book Reading and Oral Language scales), and (d) early mathematics (TBRS Math Concepts scale) in the spring of pre-kindergarten only. To analyze these data, ANCOVAs were conducted; the covariates were (a) teacher has a BA degree, (b) previous teaching experience, (c) child/adult ratio in classroom, (d) average class size, (e) city size, and (f) geographic site.

There were statistically reliable differences favoring the *Creative Curriculum* classrooms on the Written Expression ($ES_C = 1.73, p < .01$) and Oral Language ($ES_C = 1.80, p < .01$) scales. There were no statistically significant differences on the Book Reading, Print and Letter Knowledge, Phonological Awareness, or Math Concepts scales of the TBRS.

Based on the analyses of the TBRS Print and Letter Knowledge and Written Expression scales, we conclude that *Creative Curriculum* had a positive effect on early literacy instruction relative to the control condition.

Based on the analysis of the TBRS Phonological Awareness scale, we conclude that *Creative Curriculum* did not have a statistically detectable effect on instruction in phonological awareness relative to the control condition.

Based on the analysis of the TBRS Book Reading and Oral Language scales, we conclude that *Creative Curriculum* had a positive effect on language instruction relative to the control condition.

Based on the analysis of the TBRS Math Concepts scale, we conclude that *Creative Curriculum* did not have a statistically detectable effect on early mathematics instruction relative to the control condition.

Summary of Findings for *Creative Curriculum*: North Carolina and Georgia

The findings for *Creative Curriculum* are summarized in table 3.4.

Table 3.4. Effect sizes for *Creative Curriculum*: North Carolina and Georgia

Measure	Student-level effect sizes (ES_s)		
	RM analysis Spring Pre-K	RM analysis Spring K	ANCOVA Spring K
Mathematics			
WJ Applied Problems	.20	.09	—
CMA-A Mathematics Composite	-.10	.14	—
Shape Composition ¹	.19	-.01	—
Reading			
TERA	-.08	-.04	—
WJ Letter Word Identification	-.08	.00	—
WJ Spelling	-.18	-.05	—
Phonological awareness			
Pre-CTOPPP/CTOPP	.02	†	.06
Language			
PPVT	.08	.15	—
TOLD	-.16	-.17	—
Behavior			
SSRS Social Skills	.05	†	-.12
SSRS Problem Behavior ²	-.16	†	.08
PLBS/LBS	.07	†	-.20
Measure	Classroom-level effect sizes (ES_c)		
	RM analysis Spring Pre-K	ANCOVA Spring Pre-K	
Global classroom quality			
ECERS-R	1.66*	—	
Teacher-child interaction			
Arnett Detachment ³	-1.68*	—	
Arnett Harshness ³	-.70	—	
Arnett Permissiveness ³	-1.01	—	
Arnett Positive Interactions	1.65**	—	
Teacher instructional practices ⁴			
TBRS Book Reading	†	.28	
TBRS Oral Language	†	1.80**	
TBRS Phonological Awareness	†	-.10	
TBRS Print and Letter Knowledge	†	1.02	
TBRS Written Expression	†	1.73**	
TBRS Math Concepts	†	.75	

— Not available.

† Not applicable. Four of the kindergarten student-level measures were not on the same scale as the pre-kindergarten measures. The classroom-level data were only collected during the pre-kindergarten year of the study.

* $p < .05$; ** $p < .01$ ¹ Building Blocks, Shape Composition task² Higher scores on this scale represent more negative child behaviors.³ Lower scores on this scale represent a more positive classroom environment.⁴ ANCOVA models for the TBRS measures did not include baseline pretest scores because TBRS data were only collected in spring of the pre-kindergarten year.

NOTE: RM: Repeated Measures

ANCOVA: Analysis of covariance

Significance indications (p -values) in the table refer to the tests of contrasts between intervention and control groups that underlie the effect sizes reported here. Refer to the glossary for abbreviations of the measures.

SOURCE: The Preschool Curriculum Evaluation Research (PCER) Study.

Chapter 4. Creative Curriculum with Ladders to Literacy: University of New Hampshire (New Hampshire site)

Curriculum

The University of New Hampshire (New Hampshire) research team chose to evaluate *Ladders to Literacy*, an early literacy and language development supplementary curriculum for preschool and kindergarten children. The curriculum is intended for use in inclusive classrooms with children with disabilities and children for whom English is a second language. The curriculum includes more than 50 skill-building activities that are organized into three sections:

- print awareness;
- metalinguistic awareness; and
- oral language.

The activities included in the curriculum are designed to be suggestions or models that teachers can adopt for use with an existing classroom curriculum. Teachers are encouraged to select the activities that they want to implement and incorporate those activities into their daily classroom schedule. Teachers are provided with guidance on how to use scaffolding techniques to individualize children's learning of language and literacy skills.

The New Hampshire researchers selected a common subset of 27 activities that all *Ladders to Literacy* treatment group teachers used throughout the school year. For this evaluation, *Ladders to Literacy* was implemented as a supplementary curriculum to the *Creative Curriculum*. Classrooms in the control condition implemented the *Creative Curriculum* without the supplement.

Sample

The New Hampshire research team recruited Head Start classrooms to participate in the study. The Head Start program administrators and teachers received a program incentive for participating in the study. Less than half were full-day programs. A total of 14 teachers/classrooms were recruited for the study. The New Hampshire research team relied on the teachers in each of the participating classrooms to distribute consent forms to the families of eligible children. A sample of 123 children (62 treatment, 61 control) and parents were recruited to participate in the study. Data were collected on a total of 123 children and 20 parents at the time of the fall baseline data collection.

In the follow-up year of the evaluation study, the sample of schools went from 8 in pre-kindergarten to 26 schools in kindergarten. The sample of classrooms went from 14 preschool to 41 classrooms in kindergarten.

Children and Families

The children were 4.6 years old at the time of baseline data collection and less than half (44%) were male. The racial/ethnic composition of the sample of children was diverse: 39 percent White, 11 percent African American, and 31 percent identified as Hispanic. Table 4.1 provides additional information on the demographic characteristics of the children in the study sample. There were no statistically detectable differences between the treatment and control groups on these child characteristics.

The demographic characteristics of the primary caregivers, who were most often the biological or adoptive mother, are presented in table 4.2. The average age of the primary caregiver was 30 years. Less than half

Table 4.1. Child demographic characteristics for *Creative Curriculum with Ladders to Literacy*

Characteristics	Full sample n = 123	Curriculum comparison	
		Control n = 61	Treatment n = 62
Age at baseline (years), mean	4.6	4.6	4.6
Gender (% male)	43.9	41.0	46.8
Race/ethnicity (%)			
White, non-Hispanic	38.7	44.8	33.3
African American, non-Hispanic	11.3	‡	12.1
Hispanic	30.6	20.7	39.4
Asian or Pacific Islander	0.0	0.0	0.0
Native American	0.0	0.0	0.0
Multiple/other	19.4	24.1	15.2
Child disability status (parent reported, %) ¹	25.0	0.0	33.3

‡ Reporting standards not met. Values suppressed to protect participant confidentiality.

¹ Because there were so few parent interviews in New Hampshire during the baseline round, the value reported in this table might be invalid. The parent interview rates were higher at the time of the spring pre-kindergarten interview (55 parent interviews) and the spring kindergarten interview (63 parent interviews).

NOTE: Child disability status was included in the child-level impact analyses. The analysis results should be interpreted with caution because of the questionable validity of the parent report data.

SOURCE: PCER Parent Interview (Fall 2003, Spring 2004, and Spring 2005).

Table 4.2. Primary caregiver demographic characteristics for *Creative Curriculum with Ladders to Literacy*

Characteristics	Full sample n = 20 ¹	Curriculum comparison	
		Control n = 5	Treatment n = 15
Age at baseline (years), mean	30.2	31.8	29.7
Marital status (%)			
Married	40.0	40.0	40.0
Separated/Divorced	30.0	40.0	26.7
Widowed	0.0	0.0	0.0
Never Married	30.0	20.0	33.3
Race/ethnicity (%)			
White, non-Hispanic	50.0	80.0	40.0
African American, non-Hispanic	0.0	0.0	0.0
Hispanic	20.0	0.0	26.7
Asian or Pacific Islander	0.0	0.0	0.0
Native American	0.0	0.0	0.0
Multiple/other	30.0	‡	33.3
Educational level (%)			
Did not finish high school	45.0	‡	46.7
High school diploma or GED	40.0	‡	33.3
Some college	‡	0.0	‡
College graduate	‡	0.0	‡
Employment (%)			
Full-time	30.0	‡	33.3
Part-time	0.0	0.0	0.0
Unemployed	70.0	80.0	66.7
Other	0.0	0.0	0.0

‡ Reporting standards not met. Values suppressed to protect participant confidentiality.

¹ Because there were so few parent interviews in New Hampshire during the baseline round, the values reported in the tables might be invalid. The parent interview rates were higher at the time of the spring pre-kindergarten interview (55 parent interviews) and the spring kindergarten interview (63 parent interviews). Mother's educational level was included in the child-level impact analyses. The analysis results should be interpreted with caution because of the questionable validity of the parent report data.

SOURCE: PCER Parent Interview (Fall 2003, Spring 2004, and Spring 2005).

(40%) were married, 30 percent were never married, and 30 percent were separated or divorced. Few (15%) primary caregivers reported having had some college or a bachelor's degree, 40 percent had a high school diploma or GED, and 45 percent had not finished high school. About one-third (30%) of the primary caregivers were employed full-time, and 70 percent were not employed. There were no statistically detectable differences between the treatment and control groups on the primary caregiver characteristics.

Teachers

Fourteen teachers participated in the preschool year intervention study. Most of the preschool teachers were female. Most (93%) of the teachers were White. The preschool teachers had an average of 9 years of teaching experience and, on average, 8 years of experience teaching preschool. Nearly half of the teachers had either a bachelor's degree (36%) or higher, 29 percent had an associate's degree, and 29 percent had a high school diploma or GED. Forty-three percent of the teachers reported having a current teaching license or certification and 36 percent had no teacher certification credentials. Table 4.3 provides additional information on the characteristics of the preschool sample of teachers. There were no statistically detectable differences between the treatment and control groups on the teacher characteristics.

Programs/Classrooms

The average preschool class size was 13.3 children. The child-staff ratio was on average 5.8 children to one teacher or program staff person.

Random Assignment

Randomization was done during the pilot year (2002-03) of curriculum implementation, and the assignments (with some modifications) were maintained for the second year of implementation (2003-04). The research team used a random number software program to assign classrooms to conditions. In 2002, the New Hampshire research team selected 12 classrooms/teachers from a list of prospective study participants. They randomly selected four urban full-day classrooms from the list and randomly assigned two of those classrooms to the treatment condition and two to the control condition. The research team purposefully selected two urban half-day classrooms as a matched pair, with similarly high numbers of Spanish-speaking children who were enrolled in these classrooms. These two half-day classrooms were randomly assigned to the treatment and control conditions (one classroom in each condition). The research team randomly selected two additional urban half-day classrooms from the remaining list, and randomly assigned one of these classrooms to the control condition and the other to the treatment condition. The team grouped four suburban/rural classrooms that were on the list by the existence or absence of a kindergarten program in their respective towns. The research team randomly assigned one of the "kindergarten" towns and one of the "no kindergarten" towns to the control condition, and the other two classrooms to the treatment condition. A total of four full-day urban classrooms (two treatment and two control); four half-day urban classrooms (two control and two treatment); and four part-day rural classrooms (two control and two treatment) were included in the pilot-year study sample.

In the second year of implementation, the research team was able to retain 11 of the pilot-year classrooms. Two additional classrooms were added to the sample in the second year. The researchers flipped a coin to assign these two classrooms to treatment and control conditions. All of the six teachers in the treatment classrooms in the pilot year remained in the sample in the second year. Teachers were retained in three of the six control classrooms from the pilot year. One of the six control group teachers declined to participate in year 2 and was replaced by another classroom in the same center. In another two control group classrooms, the pilot-year teachers left the programs and the replacement teachers in these two classrooms agreed to participate in the study. A total of 14 classrooms and 123 children took part in the study. Details regarding randomization procedures and changes from the pilot year to the second year are described in chapter 1 of this report.

Table 4.3. Preschool teacher characteristics for *Creative Curriculum with Ladders to Literacy*

Characteristics	Full sample n = 14	Curriculum comparison	
		Control n = 7	Treatment n = 7
Gender (% female)	93.0	100.0	86.0
Race/ethnicity (%)			
White, non-Hispanic	93.0	86.0	100.0
African American, non-Hispanic	‡	‡	0.0
Hispanic	0.0	0.0	0.0
Asian or Pacific Islander	0.0	0.0	0.0
Native American	0.0	0.0	0.0
Multiple/other	0.0	0.0	0.0
Educational level (%)			
High school diploma or GED	29.0	‡	‡
Associate's degree	29.0	‡	‡
Bachelor's degree	36.0	‡	57.0
Graduate degree	7.0	‡	0.0
Current teaching license/certificate (%)	43.0	71.0	‡
Child Development Associate (CDA) (%)	‡	0.0	‡
State-awarded preschool certificate (%)	‡	0.0	‡
No credential (%)	36.0	‡	57.0
Years of teaching experience, overall (mean)	9.3	7.6	11.0
Years of preschool teaching experience (mean)	7.6	6.7	8.4

‡ Reporting standards not met. Values suppressed to protect participant confidentiality.

SOURCE: PCER Preschool Teacher Survey (Fall 2003 and Spring 2004).

Contamination

In one of the sites, there were multiple classrooms (four classrooms) of both conditions in the same building. The research team met with the treatment teachers and aides to ensure that they were not sharing any materials or activities with the control group teachers. The research team also conducted classroom observations in the treatment and control classrooms during the preschool year. There was no evidence of contamination.

Control Condition

In the control classrooms, *Creative Curriculum* was implemented as it is normally implemented (i.e., without the *Ladders to Literacy* add-on). All of the teachers received at least 1 day of *Creative Curriculum* training from a staff member at Teaching Strategies, Inc.

Data Collection

RTI International (RTI) collected the child, teacher, and school data for the New Hampshire site for all three waves of data collection. The New Hampshire research team was responsible for conducting the parent interviews in the preschool year. In the kindergarten follow-up year, RTI staff completed the parent interviews. The fall assessment data collection window for child assessments ranged from October 6, 2003 to November 24, 2003. The average delay from the beginning of the treatment (i.e., start of the school year) to the beginning of the fall assessment window was 10 days. The spring pre-kindergarten window was April 2, 2004 to June 10, 2004 and the kindergarten follow-up window was April 21, 2005 to June 24, 2005.

Attrition

Fourteen classrooms were randomly assigned to treatment or control condition. All 14 classrooms remained in the study throughout the pre-kindergarten year.

For the child assessment, the baseline (fall 2003) response rate was 100 percent; the spring 2004 response rate was 85 percent; and the kindergarten follow-up response rate was 80 percent.

Implementation

The New Hampshire research team implemented the *Ladders to Literacy* supplementary curriculum as an add-on to the existing curriculum (*Creative Curriculum*). Thus, *Creative Curriculum* was in use in both treatment and control classrooms. All of the teachers received at least 1 day of *Creative Curriculum* training from a staff member at Teaching Strategies, Inc.

The treatment group teachers received initial *Ladders to Literacy* training in September of the preschool year. The treatment group teachers received ongoing *Ladders to Literacy* training on a monthly basis throughout the school year (October 2003 to April 2004). The treatment group teachers received training to implement 27 language and literacy activities that covered three domains (print/book awareness, metalinguistic awareness, and oral language). Teachers were expected to implement nine activities (three from each of the three major domains) in the months of November and December 2003. Teachers were instructed to cumulatively add three to six additional activities on a monthly basis, from January to May 2004, following an implementation schedule that went through May of the preschool year.

Site-specific curriculum fidelity observations were conducted in all of the 14 participating classrooms on a monthly basis from December through April of the preschool year. A total of 70 observations were made during that time period. In the treatment classrooms, observations were conducted during *Ladders to Literacy* activities. In the control classrooms, observations were conducted during activities similar in form to the *Ladders to Literacy* activities (e.g., morning circle time and story reading time). A second observer conducted an observation at the same time as a primary observer on 12 occasions. The New Hampshire research team evaluated inter-observer reliability based on 17 percent of the total number of observations conducted across the preschool year. The research team observed implementation of 23 of the 27 *Ladders to Literacy* activities at least once during the year. The site-specific fidelity observations indicated that teachers in the treatment classrooms implemented more items across *Ladders to Literacy* activities than teachers in the control classrooms.

In addition to conducting *Ladders to Literacy* implementation fidelity observations, the New Hampshire research team also conducted *Creative Curriculum* fidelity observations. The research team conducted *Creative Curriculum* fidelity observations in both the treatment and control classrooms. The *Creative Curriculum* implementation fidelity checklist included items in five areas: (1) physical environment of each activity area; (2) structure (e.g., daily schedules and routines); (3) teacher-child interactions; (4) assessment (e.g., children's progress); and (5) family involvement. The New Hampshire research team collected data on physical environment, structure, and teacher-child interactions. The research team conducted two *Creative Curriculum* fidelity observations across the 14 classrooms. The mean total proportion of implementation observed was 60 percent.

Implementation Fidelity Ratings

Each research team used a global fidelity measure to rate the overall fidelity with which the curricula were implemented in the preschool year of the project. A four-point scale ranging from "Not at All" (0) to "High" (3) was used to rate each treatment classroom. Researchers were asked to use their site-specific implementation and fidelity data to rate each treatment classroom on the global fidelity measure as High, Medium, Low, or Not at All. Researchers were also asked to provide a global rating for the control group

curriculum. The *Ladders to Literacy* curriculum was rated in the high Medium range (2.71) on the global implementation fidelity measure. The control group curriculum was rated at the Medium level (2.0).

Impact Analysis Results

We present analyses for each curriculum separately, beginning with the analyses of the child-level measures (mathematics, reading, phonological awareness, and language assessments), and followed by the analyses of the classroom observation data.

Creative Curriculum with Ladders to Literacy—Child Outcomes

The unadjusted mean scores for the child-level measures are reported in table C-4a in appendix C. Covariate adjusted mean differences and standard errors are reported in table D-4a in appendix D. For all analyses of child-level measures, the following covariates were included (a) child's age, (b) gender, (c) race/ethnicity, (d) disability status as reported by parent, and (e) mother's education. The student-level effect sizes (ESs) are presented in table 4.4. There was a very low parent interview rate (20 interviews from a sample of 123) in New Hampshire during the baseline round of data collection. The parent interview rates were higher at the time of the spring pre-kindergarten interview (55 parent interviews) and the spring kindergarten interview (63 parent interviews). Child and parent background characteristics were included in the child level impact analysis models as covariates. The analysis results should be interpreted with caution because of low parent interview response rate and the questionable validity of the parent report data.

Mathematics assessments

We used repeated measures linear spline models to analyze the data from the three mathematics measures (Woodcock Johnson [WJ] Applied Problems, Child Math Assessment-Abbreviated [CMA-A] Composite Score, and Shape Composition). There were no statistically detectable differences on these measures for the fall assessment.

There were no statistically detectable differences on the spring pre-kindergarten or kindergarten assessments on any of the mathematics assessments.

Based on the analyses of the three mathematics measures, we conclude that *Ladders to Literacy* did not have a statistically detectable effect on mathematics relative to the control condition.

Reading assessments

Data from the three reading measures (Test of Early Reading Ability [TERA], WJ Letter Word Identification, and WJ Spelling) were analyzed using repeated measures linear spline models. There were no statistically detectable differences on these measures for the fall assessment.

There were no statistically detectable differences on these measures for the spring pre-kindergarten or spring kindergarten assessments.

Based on the analyses of the three reading measures, we conclude that *Ladders to Literacy* did not have a statistically detectable effect on reading relative to the control condition.

Phonological awareness

The phonological awareness measures were the Preschool Comprehensive Test of Phonological and Print Processing (Pre-CTOPPP), Elision subtest, and the Comprehensive Test of Phonological Processing (CTOPP), Kindergarten, Elision subtest. We conducted a repeated measures analysis on the Pre-CTOPPP fall and spring pre-kindergarten data. There was no statistically significant difference on the Pre-CTOPPP for the fall assessment.

There was no statistically detectable difference on the Pre-CTOPPP for the spring pre-kindergarten assessment.

We analyzed the kindergarten CTOPP data using analysis of covariance (ANCOVA). For the ANCOVA analysis, the covariates were the Pre-CTOPPP fall assessment score, child's gender, age, race/ethnicity, disability status as reported by parent, and mother's education. There was no statistically detectable difference between groups on the CTOPP for the spring kindergarten assessment.

Based on the analyses of the Pre-CTOPPP and CTOPP, we conclude that *Ladders to Literacy* did not have a statistically detectable effect on phonological awareness relative to the control condition.

Language assessments

Data from the two language measures (Peabody Picture Vocabulary Test [PPVT] and Test of Language Development [TOLD] Grammatical Understanding subtest) were analyzed using repeated measures linear spline models. There were no statistically detectable differences on these measures for the fall assessment.

There were no statistically detectable differences on these measures for the spring pre-kindergarten or spring kindergarten assessments.

Based on the analyses of the two language measures, we conclude that *Ladders to Literacy* did not have a statistically detectable effect on language development relative to the control condition.

Behavioral outcomes

We conducted a repeated measures analysis for all three pre-kindergarten social behavioral measures (Social Skills Rating System [SSRS] Social Skills scale, SSRS Problem Behaviors scale, and Preschool Learning Behavior Scale [PLBS]). The covariates were (a) child's age, (b) gender, (c) race/ethnicity, (d) disability status as reported by the parent, and (e) mother's education. There were no statistically detectable differences on these measures for the fall assessment.

For the spring pre-kindergarten assessment, there were no statistically detectable differences on any of these measures.

We analyzed the data from the kindergarten versions of the three behavioral measures (SSRS Social Skills scale, SSRS Problem Behaviors scale, and Learning Behaviors Scale [LBS]) using analysis of covariance (ANCOVA). For the ANCOVA analyses, the covariates included (a) the fall pre-kindergarten score of the pre-kindergarten version of the relevant test, along with (b) child's age, (c) gender, (d) race/ethnicity, (e) disability status as reported by the parent, and (f) mother's education.

There were no statistically detectable differences between groups on any of these measures at the spring kindergarten assessment.

Based on the analyses of the three behavioral measures, we conclude that *Ladders to Literacy* did not have a statistically detectable effect on children's social and learning behaviors relative to the control condition.

Creative Curriculum with Ladders to Literacy—Classroom Outcomes

The unadjusted mean scores for classroom measures are reported in table C-4b in appendix C. Covariate adjusted mean differences and standard errors are reported in table D-4b in appendix D. For all analyses of classroom measures, the following variables were included in the model as covariates: (a) teacher has a BA degree, (b) previous teaching experience, (c) child/adult ratio in classroom, (d) average class size, (e) city size, and (f) geographic site. The classroom-level effect sizes (E_{Sc}) are presented in table 4.4.

Overall classroom environment

We conducted a repeated measures analysis on the Early Childhood Environment Rating Scale-Revised (ECERS-R). There was no statistically detectable difference between groups on the fall observation.

No statistically detectable difference between groups was obtained for the spring pre-kindergarten observation.

Based on the analysis of the ECERS-R, we conclude that *Ladders to Literacy* did not have a statistically detectable effect on overall classroom quality relative to the control condition.

Teacher-child relationships

We obtained observations on the Arnett Detachment, Harshness, Permissiveness, and Positive Interactions scales in fall and spring of the pre-kindergarten year, and conducted repeated measures analyses. There were no statistically detectable differences on these measures for the fall observation.

There were no statistically detectable differences between groups on the Arnett Detachment, Harshness, Permissiveness, or Positive Interaction scales.

Based on the analyses of the four Arnett scales, we conclude that *Ladders to Literacy* did not have a statistically detectable effect on teacher-child relationships relative to the control condition.

Classroom instruction

We obtained observations on classroom instruction in (a) early literacy (Teacher Behavior Rating Scale [TBRS] Print and Letter Knowledge and Written Expression scales), (b) phonological awareness (TBRS Phonological Awareness scale), (c) language (TBRS Book Reading and Oral Language scales), and (d) early mathematics (TBRS Math Concepts scale) in spring of the pre-kindergarten year only. To analyze these data, ANCOVAs were conducted; the covariates were (a) teacher has a BA degree, (b) previous teaching experience, (c) child/adult ratio in classroom, (d) average class size, (e) city size, and (f) geographic site.

There was a statistically reliable difference favoring the *Creative Curriculum with Ladders to Literacy* classrooms on Written Expression ($ESc = 1.13, p < .05$). There were no statistically detectable differences on the TBRS Book Reading, Print and Letter Knowledge, Oral Language, Phonological Awareness, or Math Concepts scales.

Based on the analyses of the TBRS Print and Letter Knowledge and Written Expression scales, we conclude that *Ladders to Literacy* had a positive effect on early literacy instruction relative to the control condition.

Based on the analysis of the TBRS Phonological Awareness scale, we conclude that *Ladders to Literacy* did not have a statistically detectable effect on instruction in phonological awareness relative to the control condition.

Based on the analysis of the TBRS Book Reading and Oral Language scales, we conclude that *Ladders to Literacy* did not have a statistically detectable effect on language instruction relative to the control condition.

Based on the analysis of the TBRS Math Concepts scale, we conclude that *Ladders to Literacy* did not have a statistically detectable effect on early mathematics instruction relative to the control condition.

Summary of Findings for *Creative Curriculum with Ladders to Literacy*

The findings for *Creative Curriculum with Ladders to Literacy* are summarized in table 4.4.

Table 4.4. Effect sizes for *Creative Curriculum with Ladders to Literacy*

Measure	Student-level effect sizes (ES_s)		
	RM analysis Spring Pre-K	RM analysis Spring K	ANCOVA Spring K
Mathematics			
WJ Applied Problems	-.14	-.33	—
CMA-A Mathematics Composite	.18	-.19	—
Shape Composition ¹	.02	-.10	—
Reading			
TERA	-.30	-.54	—
WJ Letter Word Identification	-.16	-.27	—
WJ Spelling	.30	-.08	—
Phonological awareness			
Pre-CTOPPP/CTOPP	-.16	†	-.10
Language			
PPVT	-.38	-.30	—
TOLD	-.22	-.06	—
Behavior			
SSRS Social Skills	-.25	†	.17
SSRS Problem Behavior ²	-.01	†	.02
PLBS/LBS	-.08	†	-.11
	Classroom-level effect sizes (ES_c)		
Measure	RM analysis Spring Pre-K	ANCOVA Spring Pre-K	
Global classroom quality			
ECERS-R	-.71	—	
Teacher-child interaction			
Arnett Detachment ³	.51	—	
Arnett Harshness ³	-.26	—	
Arnett Permissiveness ³	1.02	—	
Arnett Positive Interactions	.03	—	
Teacher instructional practices ⁴			
TBRS Book Reading	†	-.32	
TBRS Oral Language	†	-.50	
TBRS Phonological Awareness	†	-.19	
TBRS Print and Letter Knowledge	†	.75	
TBRS Written Expression	†	1.13*	
TBRS Math Concepts	†	.44	

— Not available.

† Not applicable. Four of the kindergarten student-level measures were not on the same scale as the pre-kindergarten measures. The classroom-level data were only collected during the pre-kindergarten year of the study.

* $p < .05$ ¹ Building Blocks, Shape Composition task² Higher scores on this scale represent more negative child behaviors.³ Lower scores on this scale represent a more positive classroom environment.⁴ ANCOVA models for the TBRS measures did not include baseline pretest scores because TBRS data were only collected in spring of the pre-kindergarten year.

NOTE: RM: Repeated Measures

ANCOVA: Analysis of covariance

Significance indications (p -values) in the table refer to the tests of contrasts between intervention and control groups that underlie the effect sizes reported here. There was a very low parent interview rate (20 interviews from a sample of 123) in New Hampshire during the baseline round of data collection. The parent interview rates were higher at the time of the spring pre-kindergarten interview (55 parent interviews) and the spring kindergarten interview (63 parent interviews). Child and parent background characteristics were included in the child-level impact analysis models as covariates. The analysis results should be interpreted with caution because of the low parent interview response rate and the questionable validity of the parent report data. Refer to the glossary for abbreviations of the measures.

SOURCE: The Preschool Curriculum Evaluation Research (PCER) Study.

Page intentionally left blank.

Chapter 5. *Curiosity Corner*: Success for All Foundation (SFA sites: Florida, Kansas, and New Jersey)

Curriculum

The Success for All Foundation (SFA) research team evaluated *Curiosity Corner*, a comprehensive curriculum for 3- and 4-year-old children that was developed by SFA. The curriculum uses 38 thematic units to cover topics such as family life, opposites, seasons, and nature.

Each thematic unit includes suggested activities that are designed to promote children's language and literacy, and cognitive, mathematical, social, personal, creative, and physical development. Daily learning activities are built around learning labs, where children explore the theme through hands-on experiences and interaction with teachers. Teachers receive initial training and ongoing professional development support throughout the school year. The curriculum also features a home component, which provides families with a lending library, videos, and the opportunity for participation in classroom activities.

Sample

The SFA research team recruited preschool programs in three different states (Florida, Kansas, and New Jersey). Schools were recruited through phone calls from SFA researchers. The SFA research team targeted districts with SFA schools with preschool classes to fit their two (preschool curriculum types) x two (SFA and non-SFA kindergarten classrooms) study design. Children in the *Curiosity Corner* and control conditions in the preschool year transitioned into SFA and non-SFA schools during the kindergarten year of the study. The researchers first recruited SFA schools within a district then they asked for recommendations of non-SFA schools to participate in the study. When non-SFA schools with preschool programs were not available, the research team recruited Head Start and day care centers. The final sample included 31 teachers and classrooms. Parents were recruited with assistance from the preschool teachers, and offered an incentive to participate in the study. The average parental consent rate was 63 percent for the SFA-Florida site (61% for the treatment group, 66% for the control group); 77 percent for the SFA-Kansas site (70% for the treatment group, 83% for the control group); and 47 percent for the SFA-New Jersey site (59% for the treatment group, 38% for the control group). Across all three locations, 215 children (105 treatment, 110 control) and parents were recruited. Data were collected on a total of 211 children and 195 parents at the time of the fall assessment.

In the follow-up year of the evaluation, the sample of schools went from 18 in pre-kindergarten to 69 schools in kindergarten. The sample of classrooms went from 31 preschool to 107 kindergarten classrooms. Data were collected on 194 children and 184 parents from the original participant sample.

Children and Families

The children were 4.7 years of age at the time of baseline data collection and half (49%) were male. The sample primarily included African American (51%) and White (28%) preschoolers. The racial/ethnic composition of the sample of children varied based on the geographic location of the sample. Table 5.1 provides additional information on the demographic characteristics of the study sample. At baseline, a higher percentage of boys were in the *Curiosity Corner* classrooms relative to those assigned to the control group (61% vs. 38%, $p < .001$).

Table 5.1. Child demographic characteristics for *Curiosity Corner*

Characteristics	Full sample n = 215	Curriculum comparison	
		Control n = 110	Treatment n = 105
Age at baseline (years), mean	4.7	4.7	4.6
Gender (% male)	49.5	37.9	61.0***
Race/ethnicity (%)			
White, non-Hispanic	27.5	30.7	24.3
African American, non-Hispanic	50.5	42.6	58.3
Hispanic	13.7	20.8	6.8
Asian or Pacific Islander	2.9	0.0	5.8
Native American	‡	‡	0.0
Multiple/other	4.9	5.0	4.9
Child disability status (parent reported, %)	14.4	16.5	12.4

‡ Reporting standards not met. Values suppressed to protect participant confidentiality.

*** $p < .001$

SOURCE: PCER Parent Interview (Fall 2003, Spring 2004, and Spring 2005).

The demographic characteristics of the primary caregivers, who were most often the biological or adoptive mother, are presented in table 5.2. The average age of the primary caregiver was 31 years. Approximately one-third (35%) of the primary caregivers were not married, and 46 percent were married at the time of the baseline data collection. Half (50%) of the primary caregivers reported having had some college or a college degree, 32 percent had a high school diploma or GED, and 18 percent had not finished high school. Approximately half (51%) of the primary caregivers were employed full-time, 13 percent were employed part-time, and 36 percent were not working at the time of the fall data collection. There were no statistically detectable differences between the treatment and control groups on the primary caregiver characteristics.

Teachers

There were 31 teachers who participated in the preschool year intervention study. Almost all (97%) of the preschool teachers were female. The racial/ethnic composition of the sample included White (58%) and African American (19%) teachers. The preschool teachers had, on average, 10 years of teaching experience, and an average of 6 years of experience teaching preschool. The majority of teachers had a bachelor's (45%) or graduate (32%) degree. Sixteen percent of the teachers had an associate's degree. The majority (74%) had a preschool or regular teaching credential and 35 percent had a Child Development Associate (CDA) credential. Table 5.3 provides additional information on the characteristics of the preschool sample of teachers. There were no statistically detectable differences between the treatment and control groups on the teacher characteristics.

Programs/Classrooms

The average preschool class size was 12.7 (Kansas site), 16.8 (Florida site), and 13.7 (New Jersey site). The child-staff ratio was on average 5.5 children to one teacher or program staff person in the Kansas site, 8.4 in the Florida site, and 7.8 in the New Jersey site.

Random Assignment

SFA researchers identified school districts that had SFA and non-SFA elementary schools in the same area. The SFA research team then recruited preschool programs based on whether some of the children from each preschool would transition into both SFA and non-SFA elementary schools. For example, in a district with a

Table 5.2. Primary caregiver demographic characteristics for *Curiosity Corner*

Characteristics	Full sample n = 194	Curriculum comparison	
		Control n = 97	Treatment n = 97
Age at baseline (years), mean	31.1	31.0	31.1
Marital status (%)			
Married	46.4	42.3	50.5
Separated/Divorced	17.0	17.5	16.5
Widowed	‡	‡	‡
Never Married	35.1	39.2	30.9
Race/ethnicity (%)			
White, non-Hispanic	33.7	37.1	30.2
African American, non-Hispanic	47.2	38.1	56.3
Hispanic	14.5	21.6	7.3
Asian or Pacific Islander	2.1	‡	‡
Native American	0.0	0.0	0.0
Multiple/other	2.6	‡	‡
Educational level (%)			
Did not finish high school	17.6	16.5	18.8
High school diploma or GED	32.1	37.1	27.1
Some college	34.2	27.8	40.6
College graduate	16.1	18.6	13.5
Employment (%)			
Full-time	50.5	44.3	56.7
Part-time	13.4	14.4	12.4
Unemployed	35.6	40.2	30.9
Other	‡	‡	0.0

‡ Reporting standards not met. Values suppressed to protect participant confidentiality.

SOURCE: PCER Parent Interview (Fall 2003, Spring 2004, and Spring 2005).

preschool classroom randomly assigned to the treatment condition, it was expected that some of the children from that preschool would attend an SFA elementary school and some of the children would attend a non-SFA elementary school for kindergarten.

Along with the SFA researchers, Mathematic Policy Research, Inc. (MPR) determined the unit of random assignment for each of the three SFA research locations. The MPR research staff randomly assigned schools to curriculum conditions because a school had only one classroom or conditions dictated against varying the curriculum conditions within a school. To increase the precision with which to estimate impacts, MPR staff grouped schools into blocks of two or more and randomly assigned half the preschools in each block to the treatment group and half to the control group. MPR staff formed blocks by matching preschools on easily measured characteristics such as teachers' experience, school location, or score on a state report card system and, in doing so, ensured that those characteristics would be evenly distributed between the overall treatment and control groups. MPR staff used a random number function (RAND function in MS Excel) to generate random numbers. They sorted preschools by block and assigned a random number to each preschool. The preschools were then randomly assigned to treatment and control conditions. The staff assigned the highest-ranking preschool within the block to the treatment condition, the next highest to the control condition, alternating assignment to treatment and control conditions until all preschools were randomly assigned to one of two conditions.

Table 5.3. Preschool teacher characteristics for *Curiosity Corner*

Characteristics	Full sample n = 31	Curriculum comparison	
		Control n = 17	Treatment n = 14
Gender (% female)	97.0	94.0	100.0
Race/ethnicity (%)			
White, non-Hispanic	58.0	82.0	29.0
African American, non-Hispanic	19.0	‡	36.0
Hispanic	13.0	0.0	29.0
Asian or Pacific Islander	0.0	0.0	0.0
Native American	0.0	0.0	0.0
Multiple/other	‡	‡	‡
Educational level (%)			
High school diploma or GED	‡	‡	‡
Associate's degree	16.0	‡	‡
Bachelor's degree	45.0	53.0	36.0
Graduate degree	32.0	29.0	36.0
Current teaching license/certificate (%)	74.0	71.0	79.0
Child Development Associate (CDA) (%)	35.0	24.0	50.0
State-awarded preschool certificate (%)	‡	‡	‡
No credential (%)	‡	‡	‡
Years of teaching experience, overall (mean)	10.3	10.5	10.1
Years of preschool teaching experience (mean)	6.9	7.5	6.2

‡ Reporting standards not met. Values suppressed to protect participant confidentiality.

SOURCE: PCER Preschool Teacher Survey (Fall 2003 and Spring 2004).

Contamination

Because only one classroom from each school or preschool program participated in the evaluation, there was little risk of contamination across the treatment and control conditions.

Control Condition

In the control condition, teachers in the three geographic locations used a variety of curricula. In the Florida site, instruction was primarily based on the *Creative Curriculum* model. In Kansas, teachers used a blended curriculum including *Preschool and Language Stimulation (PALS)* and the *Animated Literacy* (Stone 2002) curriculum models, and teacher-developed curricula. In New Jersey, teachers used a teacher-developed, nonspecific curriculum.

Data Collection

MPR collected the child, parent, teacher, and school data for the SFA sites (New Jersey, Kansas, and Florida) for all three waves of data collection. The fall assessment data collection window for child assessments ranged from October 10, 2003 to November 11, 2003 (New Jersey); September 8, 2003 to November 17, 2003 (Kansas); and October 8, 2003 to November 19, 2003. The average delay from the beginning of the treatment (i.e., start of the school year) to the beginning of the fall assessment window was 35 days in New Jersey, 14 days in Kansas, and 49 days in Florida. The spring pre-kindergarten window was May 10, 2004 to June 19, 2004 (New Jersey); April 5, 2004 to May 17, 2004 (Kansas); and April 5, 2004 to May 7, 2004 (Florida). The kindergarten follow-up window was April 12, 2005 to June 8, 2005 (New Jersey); March 28, 2005 to June 8, 2005 (Kansas); and April 2, 2005 to June 8, 2005 (Florida).

Attrition

Eighteen schools were randomly assigned to treatment and control conditions. All 18 schools and 31 classrooms remained in the study throughout the pre-kindergarten year.

For the child assessment, the baseline (fall 2003) response rate was 98 percent; the spring 2004 response rate was 95 percent; and the kindergarten follow-up response rate was 90 percent.

Implementation

The *Curiosity Corner* curriculum was implemented in 14 treatment classrooms. SFA trainers provided initial training and ongoing support to teachers implementing the curriculum. This support included implementation visits, during which trainers observed teachers' instructional practices and the classroom environment. The trainers provided qualitative feedback during visits that were conducted in the fall, winter, and spring of the preschool year. The SFA curriculum fidelity instrument was used for the implementation visit report. The purpose of the fidelity measure is to determine the current level of implementation of the *Curiosity Corner* curriculum components. A team of five SFA trainers individually visited the *Curiosity Corner* classes at least three times per site beyond the initial training visits and completed implementation visit reports. During these follow-up visits, trainers provided support for teachers' emerging expertise with the program. For example, they identified areas for professional development improvement and addressed teachers' questions and concerns. They also met with administrators to discuss the results of their observations. During these meetings, strengths and areas for improvement were identified. The same team of trainers observed control classes. Using the same implementation visit rating scale used in *Curiosity Corner* classes, they visited each control classroom site at least twice to rate the classes. An initial review of the implementation visit reports and narratives indicated a wide degree of variability in the quality of implementation. Variability was evident among and within sites. The implementation quality appeared to vary by teacher. Some sites had teachers who implemented the curriculum exceptionally well and others very poorly.

Implementation Fidelity Ratings

Each research team used a global fidelity measure to rate the overall fidelity with which the curricula were implemented in the preschool year of the project. A four-point scale ranging from "Not at All" (0) to "High" (3) was used to rate each treatment classroom. Researchers were asked to use their site-specific implementation and fidelity data to rate each treatment classroom on the global fidelity measure as High, Medium, Low, or Not at All. Researchers were also asked to provide a global rating for the control group curriculum. The *Curiosity Corner* curriculum (2.0) and the control curriculum (1.9) were both rated at the Medium level (2.0) on the global implementation fidelity measure.

Impact Analysis Results

We begin with the analyses of the child-level measures (i.e., the mathematics, reading, phonological awareness, and language assessments) and then present the analyses of the classroom observation data. Our discussion of the results focuses on the combined analyses of the three SFA research sites.

Curiosity Corner—Child Outcomes

The unadjusted mean scores for child-level measures are reported in table C-5a in appendix C. Covariate adjusted mean differences and standard errors are reported in table D-5a in appendix D. For all analyses of child-level measures, the following covariates were included: (a) child's age, (b) gender, (c) race/ethnicity, (d) disability status as reported by parent, and (e) mother's education. The student-level effect sizes (ESs) are presented in table 5.4.

Mathematics assessments

We used repeated measures linear spline models to analyze the data from all three mathematics measures (Woodcock Johnson [WJ] Applied Problems, Child Math Assessment-Abbreviated [CMA-A] Composite Score, and Shape Composition). There were no statistically detectable differences between groups on these measures for the fall assessment.

There were no statistically detectable differences between groups on the spring of the pre-kindergarten or kindergarten years on any of these measures.

Based on the analyses of the three mathematics measures, we conclude that *Curiosity Corner* did not have a statistically detectable effect on mathematics relative to the control condition.

Reading assessments

Data from the three reading measures (Test of Early Reading Ability [TERA], WJ Letter Word Identification, and WJ Spelling) were analyzed using repeated measures linear spline models. There were no statistically detectable differences between groups on these measures for the fall assessment.

There was no statistically detectable difference between groups on the TERA for the spring pre-kindergarten assessment. There was, however, a statistically reliable effect for the spring kindergarten assessment ($ES_s = .43, p < .05$).

For the WJ Letter Word Identification test, there was no difference for the spring pre-kindergarten assessment, but there was a statistically reliable difference for the spring kindergarten assessment ($ES_s = .43, p < .05$).

There was no statistically detectable difference for the WJ Spelling test for either the spring pre-kindergarten or spring kindergarten assessments.

Based on the analyses of the three reading measures, we conclude that *Curiosity Corner* did not have a statistically detectable effect on pre-reading skills relative to the control condition at the end of pre-kindergarten. However, relative to the control group, results indicate there was a delayed effect of *Curiosity Corner* on reading measures at the end of the kindergarten year.

Phonological awareness

The phonological awareness measures were the Preschool Comprehensive Test of Phonological and Print Processing (Pre-CTOPPP), Elision subtest, and the Comprehensive Test of Phonological Processing (CTOPP), Kindergarten, Elision subtest. We conducted a repeated measures analysis on the Pre-CTOPPP fall and spring pre-kindergarten data. There was no statistically detectable difference on the Pre-CTOPPP for the fall assessment.

There was no statistically detectable difference on the Pre-CTOPPP for the spring pre-kindergarten assessment.

We analyzed the kindergarten CTOPP data using analysis of covariance (ANCOVA). For the ANCOVA analysis, the covariates were the (a) Pre-CTOPPP fall assessment score, (b) child's gender, (c) age, (d) race/ethnicity, (e) disability status as reported by parent, and (f) mother's education. There was no statistically detectable difference between groups on the CTOPP for the spring kindergarten assessment.

Based on the analyses of the Pre-CTOPPP and CTOPP, we conclude that *Curiosity Corner* did not have a statistically detectable effect on phonological awareness relative to the control condition.

Language assessments

Data from the two language measures (Peabody Picture Vocabulary Test [PPVT] and Test of Language Development [TOLD] Grammatical Understanding subtest) were analyzed using repeated measures linear spline models. There were no statistically detectable differences for the fall assessment on either measure.

For the spring of the pre-kindergarten and kindergarten years, there were no statistically detectable differences between groups on either measure.

Based on the analyses of the two language measures, we conclude that *Curiosity Corner* did not have a statistically detectable effect on language development relative to the control condition.

Behavioral outcomes

We conducted a repeated measures analysis for all three pre-kindergarten social behavioral measures (Social Skills Rating System [SSRS] Social Skills scale, SSRS Problem Behaviors scale, and Preschool Learning Behaviors Scale [PLBS]). The covariates were (a) child's age, (b) gender, (c) race/ethnicity, (d) disability status as reported by the parent, and (e) mother's education. There was a statistically significant difference on the SSRS Problem Behaviors scale at the fall assessment ($ES_s = .53, p < .05$); children in the *Curiosity Corner* condition were rated as exhibiting more problem behaviors relative to the control group (follow-up analyses for this finding are included in appendix A). There were no statistically significant differences on the SSRS Social Skills or PLBS for the fall assessment.

For the spring pre-kindergarten assessment, there were no statistically detectable differences on the measures of behavioral outcomes.

We analyzed the data from the kindergarten versions of the three behavioral measures (SSRS Social Skills scale, SSRS Problem Behaviors scale, and Learning Behaviors Scale [LBS]) using analysis of covariance (ANCOVA). For the ANCOVA analyses, the covariates included the fall pre-kindergarten score of the pre-kindergarten version of the relevant test, along with (a) child's age, (b) gender, (c) race/ethnicity, (d) disability status as reported by the parent, and (e) mother's education.

There were no statistically detectable differences between groups on any of these measures for the spring kindergarten assessment.

Based on the analyses of the three behavioral measures, we conclude that *Curiosity Corner* did not have a statistically detectable effect on children's social and learning behaviors relative to the control condition.

Curiosity Corner—Classroom Outcomes

The unadjusted mean scores for classroom measures are reported in table C-5b in appendix C. Covariate adjusted mean differences and standard errors are reported in table D-5b in appendix D. For all analyses of classroom measures, the following variables were included in the model as covariates: (a) teacher has a BA degree, (b) previous teaching experience, (c) child/adult ratio in classroom, (d) average class size, (e) city size, and (f) geographic site. The classroom-level effect sizes (ES_c) are presented in table 5.4.

Overall classroom environment

We conducted a repeated measures analysis on the Early Childhood Environment Rating Scale-Revised (ECERS-R). There was no statistically significant difference between groups on the fall observation.

No statistically detectable difference between groups was obtained in spring of the pre-kindergarten year.

Based on the analysis of the ECERS-R, we conclude that *Curiosity Corner* did not have a statistically detectable effect on overall classroom quality relative to the control condition.

Teacher-child relationships

We obtained observations on the Arnett Detachment, Harshness, Permissiveness, and Positive Interactions scales in fall and spring of the pre-kindergarten year, and conducted repeated measures analyses. There was a statistically significant difference at the time of the fall observation on the Arnett Permissiveness scale ($ES_c = -1.46, p < .05$; follow-up analyses for this finding are included in appendix A). Teachers in the *Curiosity Corner* classrooms were rated as being less permissive in their interactions with their students relative to teachers in the control classrooms. (Please see appendix A for additional analyses.) There were no statistically detectable differences on the other scales in fall of the pre-kindergarten year.

For the spring pre-kindergarten assessment, there were no statistically detectable differences between groups on the Arnett Detachment, Harshness, Permissiveness, or Positive Interaction scales.

Based on the analyses of the four Arnett scales, we conclude that *Curiosity Corner* did not have a statistically detectable effect on teacher-child relationships relative to the control condition.

Classroom instruction

We obtained observations on classroom instruction in (a) early literacy (Teacher Behavior Rating Scale [TBRS] Print and Letter Knowledge and Written Expression scales), (b) phonological awareness (TBRS Phonological Awareness scale), (c) language (TBRS Book Reading and Oral Language scales), and (d) early mathematics (TBRS Math Concepts scale) in the spring of pre-kindergarten only. To analyze these data, ANCOVAs were conducted; the covariates were: (a) teacher has a BA degree, (b) previous teaching experience, (c) child/adult ratio in classroom, (d) average class size, (e) city size, and (f) geographic site.

There were no statistically detectable differences on the Print and Letter Knowledge, Written Expression, Phonological Awareness, Oral Language, or Math Concepts scales. There was a statistically significant difference on the TBRS Book Reading scale ($ES_C = 2.06, p < .001$) indicating that the *Curiosity Corner* teachers provided more book reading activities relative to teachers in the control classrooms.

Based on the analyses of the TBRS Print and Letter Knowledge and Written Expression scales, we conclude that *Curiosity Corner* did not have a statistically detectable effect on early literacy instruction relative to the control condition.

Based on the analysis of the TBRS Phonological Awareness scale, we conclude that *Curiosity Corner* did not have a statistically detectable effect on instruction in phonological awareness relative to the control condition.

Based on the analysis of the TBRS Book Reading and Oral Language scales, we conclude that *Curiosity Corner* had a positive effect on language instruction relative to the control condition.

Based on the analysis of the TBRS Math Concepts scale, we conclude that *Curiosity Corner* did not have a statistically detectable effect on early mathematics instruction relative to the control condition.

Summary of Findings for *Curiosity Corner*

The findings for *Curiosity Corner* are summarized in table 5.4.

Table 5.4. Effect sizes for *Curiosity Corner*

Measure	Student-level effect sizes (ES _s)		
	RM analysis Spring Pre-K	RM analysis Spring K	ANCOVA Spring K
Mathematics			
WJ Applied Problems	.10	.26	—
CMA-A Mathematics Composite	.01	-.05	—
Shape Composition ¹	.16	.32	—
Reading			
TERA	.10	.43*	—
WJ Letter Word Identification	.09	.43*	—
WJ Spelling	.04	.20	—
Phonological awareness			
Pre-CTOPPP/CTOPP	.18	†	.25
Language			
PPVT	-.01	.14	—
TOLD	-.08	.15	—
Behavior			
SSRS Social Skills	-.06	†	.32
SSRS Problem Behavior ²	.43	†	-.08
PLBS/LBS	-.25	†	.11
Measure	Classroom-level effect sizes (ES _c)		
	RM analysis Spring Pre-K	ANCOVA Spring Pre-K	
Global classroom quality			
ECERS-R	-.48	—	
Teacher-child interaction			
Arnett Detachment ³	-.41	—	
Arnett Harshness ³	.14	—	
Arnett Permissiveness ³	-.98	—	
Arnett Positive Interactions	.02	—	
Teacher instructional practices ⁴			
TBRS Book Reading	†	2.06***	
TBRS Oral Language	†	.37	
TBRS Phonological Awareness	†	.44	
TBRS Print and Letter Knowledge	†	-.99	
TBRS Written Expression	†	-.54	
TBRS Math Concepts	†	-.33	

— Not available.

† Not applicable. Four of the kindergarten student-level measures were not on the same scale as the pre-kindergarten measures. The classroom-level data were only collected during the pre-kindergarten year of the study.

* $p < .05$; *** $p < .001$ ¹ Building Blocks, Shape Composition task² Higher scores on this scale represent more negative child behaviors.³ Lower scores on this scale represent a more positive classroom environment.⁴ ANCOVA models for the TBRS measures did not include baseline pretest scores because TBRS data were only collected in spring of the pre-kindergarten year.

NOTE: RM: Repeated Measures

ANCOVA: Analysis of covariance

Significance indications (p -values) in the table refer to the tests of contrasts between intervention and control groups that underlie the effect sizes reported here. Refer to the glossary for abbreviations of the measures.

SOURCE: The Preschool Curriculum Evaluation Research (PCER) Study.

Page intentionally left blank.

Chapter 6. *Doors to Discovery* and *Let's Begin with the Letter People*: University of Texas Health Science Center at Houston (Texas site)

Curriculum

The University of Texas Health Science Center at Houston (Texas) researchers implemented the *Doors to Discovery* and *Let's Begin with the Letter People* curricula.

Doors to Discovery

The *Doors to Discovery* curriculum is a pre-kindergarten program that is based on the five areas identified by the International Reading Association and the National Association for the Education of Young Children as the foundation for early literacy success: oral language, phonological awareness, concepts of print, alphabet knowledge and writing, and comprehension.

The program focuses on the use of learning centers and shared literacy activities in the pre-kindergarten classroom. The curriculum is presented in eight thematic units that cover topics such as friendship, communities, nature, society, and health. Classroom practices include teachers' directed activities; large and small group activities; and children's application of skills and independent practice on activities that are tied to the curriculum.

The curriculum components also include family learning activities that are designed to foster partnerships between the school and the family; initial training for teachers and ongoing professional development support; and assessment strategies that are integrated into the curriculum units.

Let's Begin with the Letter People

Like *Doors to Discovery*, *Let's Begin with the Letter People* is a comprehensive pre-kindergarten curriculum that is organized thematically. Literacy learning is integrated across topic areas including science, health and safety, art, mathematics, spatial concepts, and music, as well as development of large and small motor skills. The curriculum focuses on specific literacy and language skills including oral language, phonological and phonemic awareness, and letter knowledge. The curriculum lessons address the development of letter knowledge in multiple contexts (e.g., circle time, small group, large group) and activities (e.g., center activities, story times) that support children's development of language and literacy skills.

The teacher lesson plans incorporate activities from the thematic units that are consistent with the overall *Letter People* curriculum objectives. Classroom practices include teacher directed activities, application of skills, and independent practice on activities that are tied to the curriculum. The physical layout of the *Letter People* classroom includes clearly defined interest centers (e.g., Paint Corner, Block, Drama Center, Mathematics, etc.). The curriculum materials include Letter People (huggables). Each Letter Person represents a letter of the alphabet, and has distinguishing characteristics that is readily associated with the sound represented by that letter.

Sample

The Texas research team recruited Head Start and public pre-kindergarten (Title I and non-Title I) programs for participation in the study. All of the programs were full-day programs. All schools and teachers were recruited before the start of the preschool year. Parental consent was obtained during the first few weeks of the school year. A total of 95 teachers and 625 parents and children were recruited as part of the site-specific study. A subset of 44 teachers/classrooms, and 297 parents and children (101 in *Doors to Discovery* treatment

group, 100 in the *Let's Begin with the Letter People* treatment group, and 96 in the control group) were selected for inclusion in the study sample for the PCER initiative. Data were collected on 293 children and 237 parents at the time of the fall assessment data collection.

In the follow-up year of the evaluation, the sample of schools went from 19 in preschool to 78 schools in kindergarten. The sample of classrooms went from 44 preschool to 149 kindergarten classrooms. The kindergarten sample included 250 children and 264 parents from the original sample of 297 participants. Data were collected on 235 children and 203 parents.

Children and Families

The children were 4.6 years of age at the time of baseline data collection and more than half (55%) were male. The racial/ethnic composition of the sample of children was diverse: 43 percent Hispanic, 30 percent White, and 13 percent African American. Table 6.1 provides additional information on the demographic characteristics of the children in the study sample. There were no statistically detectable differences between the treatment and control groups on these child characteristics.

Table 6.1. Child demographic characteristics for *Doors to Discovery* and *Let's Begin with the Letter People*

Characteristics	Full sample n = 297	Curriculum comparison		
		Control n = 96	Treatment 1 ¹ n = 100	Treatment 2 ² n = 101
Age at baseline (years), mean	4.6	4.7	4.6	4.7
Gender (% male)	54.6	54.3	55.6	54.0
Race/ethnicity (%)				
White, non-Hispanic	30.1	24.1	33.3	32.9
African American, non-Hispanic	13.3	19.3	7.4	12.9
Hispanic	43.0	39.8	49.4	40.0
Asian or Pacific Islander	4.4	7.2	3.7	‡
Native American	0.0	0.0	0.0	0.0
Multiple/other	9.2	9.6	6.2	11.8
Child disability status (parent reported, %)	12.3	16.7	10.1	10.7

‡ Reporting standards not met. Values suppressed to protect participant confidentiality.

¹In Texas, Treatment 1 is *Let's Begin with the Letter People*.

²In Texas, Treatment 2 is *Doors to Discovery*.

SOURCE: PCER Parent Interview (Fall 2003, Spring 2004, and Spring 2005).

The demographic characteristics of the primary caregivers, who were most often the biological or adoptive mother, are presented in table 6.2. The average age of the primary caregiver was 34 years. The majority (72%) of the primary caregivers were married. Most reported having had some college (27%) or a bachelor's degree (34%), 17 percent had a high school diploma or GED, and 22 percent had not finished high school. Less than half (40%) of the primary caregivers were employed full-time; 20 percent were employed part-time; and 39 percent were unemployed. There were no statistically detectable differences between the treatment and control groups on the primary caregiver characteristics.

Table 6.2. Primary caregiver demographic characteristics for *Doors to Discovery* and *Let's Begin with the Letter People*

Characteristics	Full sample n = 237	Curriculum comparison		
		Control n = 73	Treatment 1 ¹ n = 79	Treatment 2 ² n = 85
Age at baseline (years), mean	34.2	34.6	34.0	34.0
Marital status (%)				
Married	71.7	64.4	75.9	74.1
Separated/Divorced	11.8	13.7	11.4	10.6
Widowed	‡	‡	0.0	‡
Never Married	15.2	20.5	12.7	12.9
Race/ethnicity (%)				
White, non-Hispanic	29.9	21.1	29.1	38.1
African American, non-Hispanic	12.8	18.3	8.9	11.9
Hispanic	43.2	45.1	48.1	36.9
Asian or Pacific Islander	4.7	5.6	5.1	‡
Native American	0.0	0.0	0.0	0.0
Multiple/other	9.4	9.9	8.9	9.5
Educational level (%)				
Did not finish high school	21.8	22.5	19.0	23.8
High school diploma or GED	16.7	15.5	21.5	13.1
Some college	27.4	36.6	25.3	21.4
College graduate	34.2	25.4	34.2	41.7
Employment (%)				
Full-time	39.7	45.2	40.5	34.1
Part-time	19.8	16.4	19.0	23.5
Unemployed	39.2	38.4	36.7	42.4
Other	‡	0.0	‡	0.0

‡ Reporting standards not met. Values suppressed to protect participant confidentiality.

¹In Texas, Treatment 1 is *Let's Begin with the Letter People*.

²In Texas, Treatment 2 is *Doors to Discovery*.

SOURCE: PCER Parent Interview (Fall 2003, Spring 2004, and Spring 2005).

Teachers

There were 44 teachers who participated in the preschool year intervention study. Most (43 of 44) of the preschool teachers were female, and most were White (55%) or African American (32%). The preschool teachers had on average 14 years of teaching experience, with an average of 8 years of experience teaching preschool. Most of the teachers had a bachelor's (66%) or graduate (14%) degree. Eleven percent of the teachers had an associate's degree, and 9 percent had a high school diploma or GED. The teachers reported having a state-awarded preschool certificate (74%), teaching license or certificate (73%), or a Child Development Associate (CDA) credential (18%). Table 6.3 provides additional information on the characteristics of the preschool sample of teachers. At baseline, the *Doors to Discovery* treatment group teachers had more years of preschool teaching experience relative to the teachers assigned to the *Let's Begin with the Letter People* and control group conditions (10.1 years vs. 8.5 years and 5.8 years, $p < .05$).

Programs/Classrooms

The average preschool class size was 18.6 children. The child-staff ratio was on average 9.1 children to one teacher or program staff person.

Table 6.3. Preschool teacher characteristics for *Doors to Discovery* and *Let's Begin with the Letter People*

Characteristics	Full sample n = 44	Curriculum comparison		
		Control n = 16	Treatment 1 ¹ n = 15	Treatment 2 ² n = 13
Gender (% female)	98.0	100.0	100.0	92.0
Race/ethnicity (%)				
White, non-Hispanic	55.0	50.0	60.0	54.0
African American, non-Hispanic	32.0	38.0	27.0	31.0
Hispanic	‡	‡	‡	‡
Asian or Pacific Islander	‡	‡	0.0	0.0
Native American	0.0	0.0	0.0	0.0
Multiple/other	‡	0.0	‡	‡
Educational level (%)				
High school diploma or GED	9.0	0.0	‡	‡
Associate's degree	11.0	13.0	‡	‡
Bachelor's degree	66.0	69.0	73.0	54.0
Graduate degree	14.0	‡	‡	‡
Current teaching license/certificate (%)	73.0	81.0	67.0	69.0
Child Development Associate (CDA) (%)	18.0	‡	‡	‡
State-awarded preschool certificate (%)	74.0	73.0	73.0	77.0
No credential (%)	‡	‡	0.0	‡
Years of teaching experience, overall (mean)	14.1	11.9	15.2	15.4
Years of preschool teaching experience (mean)	8.0	5.8	8.5	10.1*

‡ Reporting standards not met. Values suppressed to protect participant confidentiality.

* $p < .05$

¹ In Texas, Treatment 1 is *Let's Begin with the Letter People*.

² In Texas, Treatment 2 is *Doors to Discovery*.

SOURCE: PCER Preschool Teacher Survey (Fall 2003 and Spring 2004).

Random Assignment

Randomization was done during the pilot year of curriculum implementation, and the same assignments were maintained for the second year (2003-04) of implementation as had been used during the pilot study year (2002-03). Most of the teachers were second-year implementers of the *Doors to Discovery* or *Let's Begin with the Letter People* curricula. There were no changes to the group of Head Start teachers and classrooms during the second year of implementation. In the Title I district, one *Let's Begin with the Letter People* teacher was replaced; one *Doors to Discovery* teacher was replaced; and three control group teachers were replaced. In the non-Title I district, one *Let's Begin with the Letter People* teacher was replaced and one control group teacher was replaced.

The Texas research team randomly assigned 76 classrooms using a 3 (Type of Curriculum—*Let's Begin with the Letter People*, *Doors to Discovery*, or Control) x 2 (mentoring versus nonmentoring) design with classrooms from three settings (Head Start, Title I pre-kindergarten, and non-Title I pre-kindergarten). There were 27 control classrooms; 24 *Let's Begin with the Letter People* classrooms; and 25 *Doors to Discovery* classrooms dispersed across the three types of preschool settings. A subset of 45 classrooms was randomly selected for inclusion in the study for the PCER initiative. The 76 preschool teachers were provided with a description of the study and given the option to participate. The names of teachers who consented to participate were included in a hat and 45 classrooms/teachers were randomly selected. One teacher later decided not to participate, and this teacher/classroom was dropped from the sample. The final sample included 44 preschool classrooms. Eight children per classroom were randomly selected for pre- and post-testing from among the larger pool of consented children.

The research team randomly assigned preschools to conditions by first creating a list of all of the preschools and the number of available classrooms in each preschool building. The team then labeled three chips with the names of one of the three conditions (*Doors to Discovery*, *Let's Begin with the Letter People*, and Control) and placed the chips in a box. A chip was randomly pulled from the box and the preschool (and related classrooms) on the list was assigned to that condition. For example, school number three on the list was assigned to the control condition if that chip was pulled from the box when the team got to school number three on their list of schools. This procedure was repeated for different school sites until the target number of classrooms was obtained. All of the preschool classrooms at a preschool site/building were assigned to the same condition. The same procedure that was used to assign schools and classrooms to a condition was used to assign treatment classrooms to the mentoring (mentoring vs. nonmentoring) conditions. The names of teachers in a given treatment condition, within a type of pre-kindergarten setting (e.g., 10 Head Start *Let's Begin with the Letter People* classrooms), were put into a container. Half of the teachers were randomly selected to receive mentoring along with their implementation of the treatment curriculum.

Across the three types of preschool settings, a total of 15 classrooms received training and implemented the *Let's Begin with the Letter People* curriculum and 14 classrooms received training and implemented the *Doors to Discovery* curriculum. Half of the teachers in each treatment curriculum condition were randomly assigned to receive mentoring support on a weekly basis. Twenty-seven classrooms were randomly assigned to a control condition that received no specific curriculum, training, or mentoring. A total of 44 classrooms and 297 children took part in the study.

Contamination

Because all classrooms at a preschool site were assigned to only one of three conditions, there was little risk of contamination across the two treatment and control conditions.

Control Condition

In classrooms in the control condition, teachers used teacher-developed, nonspecific curricula.

Data Collection

RTI International (RTI) collected the child, teacher, and school data for the Texas site for all three waves of data collection. The Texas research team was responsible for conducting the parent interviews in the preschool year. In the kindergarten follow-up year, RTI staff completed the parent interviews. The fall assessment data collection window for child assessments ranged from September 8, 2003 to October 29, 2003. The average delay from the beginning of the treatment (i.e., start of the school year) to the beginning of the fall assessment window was 20 days. The spring pre-kindergarten window was April 15, 2004 to June 11, 2004, and the kindergarten follow-up window was April 8, 2005 to June 29, 2005.

Attrition

Forty-four classrooms were randomly assigned to treatment and control conditions. All 44 classrooms remained in the study throughout the pre-kindergarten year.

For the child assessment, the baseline (fall, 2003) response rate was 99 percent, the spring 2004 response rate was 94 percent, and the kindergarten follow-up response rate was 94 percent.

Implementation

Teachers received curriculum implementation training prior to the start of the 2003-04 school year. The teacher sample included 45 teachers who participated in the pilot year of the study (2002-03), and seven new teachers who started in 2003-04. A total of 44 (37 returning) teachers participated in the study during the second year of implementation. The new teachers received 12 hours and returning teachers received 6 hours of curriculum implementation training.

The research team collected site-specific curriculum fidelity data three times during the preschool year. Control classrooms were not observed using the curriculum-specific fidelity measures. All classrooms were observed using a site-specific measure (the Teacher Behavior Rating Scale) in both treatment and control classrooms in fall and spring of the preschool year.

Implementation Fidelity Ratings

Doors to Discovery

Each research team used a global fidelity measure to rate the overall fidelity with which the curricula were implemented in the preschool year of the project. A four-point scale ranging from “Not at All” (0) to “High” (3) was used to rate each treatment classroom. Researchers were asked to use their site-specific implementation and fidelity data to rate each treatment classroom on the global fidelity measure as High, Medium, Low, or Not at All. Researchers were also asked to provide a global rating for the control group curriculum. The *Doors to Discovery* curriculum was rated Medium (2.13) on the global fidelity measure. The control group classrooms were rated as Low (1.0).

Let's Begin with the Letter People

Each research team used a global fidelity measure to rate the overall fidelity with which the curricula were implemented in the preschool year of the project. A four-point scale ranging from “Not at All” (0) to “High” (3) was used to rate each treatment classroom. Researchers were asked to use their site-specific implementation and fidelity data to rate each treatment classroom on the global fidelity measure as High, Medium, Low, or Not at All. Researchers were also asked to provide a global rating for the control group curriculum. The *Let's Begin with the Letter People* curriculum was rated Medium (1.86) on the global implementation fidelity measure. The control group curriculum was rated Low (1.0) on implementation fidelity level.

Impact Analysis Results

We present analyses for each curriculum separately beginning with the analyses of the child-level measures (i.e., the mathematics, reading, phonological awareness, and language assessments) followed by the analyses of the classroom observation data. We first present from the *Doors to Discovery* analyses and then from the *Let's Begin with the Letter People* analyses.

Doors to Discovery—Child Outcomes

The unadjusted mean scores for child-level measures are reported in table C-6a in appendix C. Covariate adjusted mean differences and standard errors are reported in table D-6a in appendix D. For all analyses of child-level measures, the following covariates were included: (a) child's age, (b) gender, (c) race/ethnicity, (d) disability status as reported by parent, and (e) mother's education. The student-level effect sizes (ESs) are presented in table 6.4.

Mathematics assessments

We used repeated measures linear spline models to analyze the data from all three mathematics measures (Woodcock Johnson [WJ] Applied Problems, Child Math Assessment-Abbreviated [CMA-A] Composite Score, and Shape Composition). There were no statistically detectable differences between the *Doors to Discovery* group and the control group on any of these measures for the fall assessment.

There were no statistically detectable differences on any of these measures for the spring pre-kindergarten or spring kindergarten assessments.

Based on the analyses of the three mathematics measures, we conclude that *Doors to Discovery* did not have a statistically detectable effect on mathematics relative to the control condition.

Reading assessments

Data from the three reading measures (Test of Early Reading Ability [TERA], WJ Letter Word Identification, and WJ Spelling) were analyzed using repeated measures linear spline models. There were no statistically detectable differences on any of these measures at the fall assessment.

There were no statistically detectable differences on any of these measures for the spring pre-kindergarten or spring kindergarten assessments.

Based on the analyses of the three reading measures, we conclude that *Doors to Discovery* did not have a statistically detectable effect on reading relative to the control condition.

Phonological awareness

The phonological awareness measures were the Preschool Comprehensive Test of Phonological and Print Processing (Pre-CTOPPP), Elision subtest, and the Comprehensive Test of Phonological Processing (CTOPP), Kindergarten, Elision subtest. We conducted a repeated measures analysis on the Pre-CTOPPP fall and spring pre-kindergarten data. There was no statistically detectable difference on the Pre-CTOPPP for the fall assessment.

There was no statistically detectable difference on the Pre-CTOPPP for the spring pre-kindergarten assessment.

We analyzed the kindergarten CTOPP data using analysis of covariance (ANCOVA). For the ANCOVA analysis, the covariates were the (a) Pre-CTOPPP fall assessment score, (b) child's gender, (c) age, (d) race/ethnicity, (e) disability status as reported by parent, and (f) mother's education. There was no statistically detectable difference between groups on the CTOPP for the spring kindergarten assessment.

Based on the analyses of the Pre-CTOPPP and CTOPP, we conclude that *Doors to Discovery* did not have a statistically detectable effect on phonological awareness relative to the control condition.

Language assessments

Data from the two language measures (Peabody Picture Vocabulary Test [PPVT] and Test of Language Development [TOLD] Grammatical Understanding subtest) were analyzed using repeated measures linear spline models. At the fall assessment, there was no statistically detectable difference between groups on the PPVT, but there was a statistically reliable difference favoring the *Doors to Discovery* group on the TOLD Grammatical Understanding scale ($ES_s = .38, p < .05$; follow-up analyses for this finding are included in appendix A).

There were no statistically detectable differences on either of these measures for the spring pre-kindergarten or spring kindergarten assessments. Based on the analyses of the two language measures, we conclude that *Doors to Discovery* did not have a statistically detectable effect on language development relative to the control condition.

Behavioral outcomes

We conducted a repeated measures analysis for all three pre-kindergarten social behavioral measures (Social Skills Rating System [SSRS] Social Skills scale, SSRS Problem Behaviors scale, and Preschool Learning Behaviors Scale [PLBS]). The covariates were (a) child's age, (b) gender, (c) race/ethnicity, (d) disability status as reported by the parent, and (e) mother's education. There were no statistically detectable differences on these measures for the fall assessment.

For the spring pre-kindergarten assessment, there were no statistically detectable differences on any of these measures.

We analyzed the data from the kindergarten versions of the three behavioral measures (SSRS Social Skills scale, SSRS Problem Behaviors scale, and Learning Behaviors Scale [LBS]) using analysis of covariance (ANCOVA). For the ANCOVA analyses, the covariates included (a) the fall pre-kindergarten score of the pre-kindergarten version of the relevant test, along with (b) child's age, (c) gender, (d) race/ethnicity, (e) disability status as reported by the parent, and (f) mother's education.

There were no statistically detectable differences between groups on the three behavioral measures for the spring kindergarten assessment.

Based on the analyses of the three behavioral measures, we conclude that *Doors to Discovery* did not have a statistically detectable effect on children's social and learning behaviors relative to the control condition.

Doors to Discovery—Classroom Outcomes

The unadjusted mean scores for classroom measures are reported in table C-6b in appendix C. Covariate adjusted mean differences and standard errors are reported in table D-6b in appendix D. For all analyses of classroom measures, the following variables were included in the model as covariates: (a) teacher has a BA degree, (b) previous teaching experience, (c) child/adult ratio in classroom, (d) average class size, (e) city size, and (f) geographic site. The classroom-level effect sizes (ES_c) are presented in table 6.4.

Overall classroom environment

We conducted a repeated measures analysis on the Early Childhood Environment Rating Scale-Revised (ECERS-R). There was no statistically detectable difference between groups on the fall observation.

No statistically detectable difference between groups was obtained for the spring pre-kindergarten observation.

Based on the analysis of the ECERS-R, we conclude that *Doors to Discovery* did not have a statistically detectable effect on overall classroom quality relative to the control condition.

Teacher-child relationships

We obtained observations on the Arnett Detachment, Harshness, Permissiveness, and Positive Interactions scales in fall and spring of the pre-kindergarten year, and conducted repeated measures analyses. There was a statistically significant difference at the time of the fall observation on the Arnett Permissiveness scale ($ES_c = 1.06, p < .05$; follow-up analyses for this finding are included in appendix A). Teachers in the *Doors to Discovery* classrooms were rated as being more permissive in their interactions with their students relative to teachers in the control classrooms. There were no statistically detectable differences on the other scales for the fall observation.

There were no statistically detectable differences between groups on the Arnett Detachment, Harshness, Permissiveness, or Positive Interactions scales for the spring pre-kindergarten assessment.

Based on the analyses of the four Arnett scales, we conclude that *Doors to Discovery* did not have a statistically detectable effect on teacher-child relationships relative to the control condition.

Classroom instruction

We obtained observations on classroom instruction in (a) early literacy (Teacher Behavior Rating Scale [TBRS] Print and Letter Knowledge and Written Expression scales), (b) phonological awareness (TBRS Phonological Awareness scale), (c) language (TBRS Book Reading and Oral Language scales), and (d) early mathematics (TBRS Math Concepts scale) in spring of pre-kindergarten only. To analyze these data, ANCOVAs were conducted; the covariates were: (a) teacher has a BA degree, (b) previous teaching experience, (c) child/adult ratio in classroom, (d) average class size, (e) city size, and (f) geographic site.

There were no statistically detectable differences on the Written Expression, Phonological Awareness, Oral Language, or Math Concepts scales. There were statistically significant differences on the Book Reading ($ES_c = 1.18, p < .01$) and Print and Letter Knowledge ($ES_c = .90, p < .05$) scales indicating that the *Doors to Discovery* teachers provided more instruction in Book Reading and Print and Letter Knowledge relative to teachers in the control classrooms.

Based on the analyses of the TBRS Print and Letter Knowledge and Written Expression scales, we conclude that *Doors to Discovery* had a positive effect on early literacy instruction relative to the control condition.

Based on the analysis of the TBRS Phonological Awareness scale, we conclude that *Doors to Discovery* did not have a statistically detectable effect on instruction in phonological awareness relative to the control condition.

Based on the analysis of the TBRS Book Reading and Oral Language scales, we conclude that *Doors to Discovery* had a positive effect on language instruction relative to the control condition.

Based on the analysis of the TBRS Math Concepts scale, we conclude that *Doors to Discovery* did not have a statistically detectable effect on early mathematics instruction relative to the control condition.

Summary of Findings for *Doors to Discovery*

The findings for *Doors to Discovery* are summarized in table 6.4 .

Table 6.4. Effect sizes for *Doors to Discovery*

Measure	Student-level effect sizes (ES _s)		
	RM analysis Spring Pre-K	RM analysis Spring K	ANCOVA Spring K
Mathematics			
WJ Applied Problems	.01	-.02	—
CMA-A Mathematics Composite	.13	-.16	—
Shape Composition ¹	-.13	-.12	—
Reading			
TERA	.06	-.05	—
WJ Letter Word Identification	.10	-.09	—
WJ Spelling	.06	-.12	—
Phonological awareness			
Pre-CTOPPP/CTOPP	.18	†	-.09
Language			
PPVT	.15	.18	—
TOLD	.17	.06	—
Behavior			
SSRS Social Skills	-.18	†	-.05
SSRS Problem Behavior ²	-.14	†	.46
PLBS/LBS	-.18	†	-.32
Measure	Classroom-level effect sizes (ES _c)		
	RM analysis Spring Pre-K	ANCOVA Spring Pre-K	
Global classroom quality			
ECERS-R	.39	—	
Teacher-child interaction			
Arnett Detachment ³	-.07	—	
Arnett Harshness ³	-.38	—	
Arnett Permissiveness ³	.13	—	
Arnett Positive Interactions	.38	—	
Teacher instructional practices ⁴			
TBRS Book Reading	†	1.18**	
TBRS Oral Language	†	.59	
TBRS Phonological Awareness	†	.58	
TBRS Print and Letter Knowledge	†	.90*	
TBRS Written Expression	†	.62	
TBRS Math Concepts	†	.37	

— Not available.

† Not applicable. Four of the kindergarten student-level measures were not on the same scale as the pre-kindergarten measures. The classroom-level data were only collected during the pre-kindergarten year of the study.

* $p < .05$; ** $p < .01$

¹ Building Blocks, Shape Composition task

² Higher scores on this scale represent more negative child behaviors.

³ Lower scores on this scale represent a more positive classroom environment.

⁴ ANCOVA models for the TBRS measures did not include baseline pretest scores because TBRS data were only collected in spring of the pre-kindergarten year.

NOTE: RM: Repeated Measures

ANCOVA: Analysis of covariance

Significance indications (p -values) in the table refer to the tests of contrasts between intervention and control groups that underlie the effect sizes reported here. Refer to the glossary for abbreviations of the measures.

SOURCE: The Preschool Curriculum Evaluation Research (PCER) Study.

Let's Begin with the Letter People—Child Outcomes

The unadjusted mean scores for child-level measures are reported in table C-7a in appendix C. Covariate adjusted mean differences and standard errors are reported in table D-7a in appendix D. For all analyses of child-level measures, the following covariates were included (a) child's age, (b) gender, (c) race/ethnicity, (d) disability status as reported by parent, and (e) mother's education. The student-level effect sizes (ESs) are presented in table 6.5.

Mathematics assessments

We used repeated measures linear spline models to analyze the data from all three mathematics measures (WJ Applied Problems, CMA-A Composite Score, and Shape Composition). There were no statistically detectable differences on any of these measures at the fall assessment.

There were no statistically detectable differences between groups on any of these measures for the spring pre-kindergarten or spring kindergarten assessments.

Based on analyses of the three mathematics measures, we conclude that *Let's Begin with the Letter People* did not have a statistically detectable effect on mathematics relative to the control condition.

Reading assessments

Data from the three reading measures (TERA, WJ Letter Word Identification, and WJ Spelling) were analyzed using repeated measures linear spline models. There were no statistically detectable differences on any of these measures at the fall assessment.

There were no statistically detectable differences between groups on any of these measures for the spring pre-kindergarten or spring kindergarten assessments.

Based on analyses of the three reading measures, we conclude that *Let's Begin with the Letter People* did not have a statistically detectable effect on reading relative to the control condition.

Phonological awareness

The phonological awareness measures were the Preschool Comprehensive Test of Phonological and Print Processing (Pre-CTOPPP), Elision subtest, and the Comprehensive Test of Phonological Processing (CTOPP), Kindergarten, Elision subtest. We conducted a repeated measures analysis on the Pre-CTOPPP fall and spring pre-kindergarten data. There was no statistically detectable difference on the Pre-CTOPPP for the fall assessment.

There was no statistically detectable difference on the Pre-CTOPPP for the spring pre-kindergarten assessment.

We analyzed the kindergarten CTOPP data using analysis of covariance (ANCOVA). For the ANCOVA analysis, the covariates were the (a) Pre-CTOPPP fall assessment score, (b) child's gender, (c) age, (d) race/ethnicity, (e) disability status as reported by parent, and (f) mother's education. There was no statistically significant difference between groups on the CTOPP for the spring kindergarten assessment.

Based on the analyses of the Pre-CTOPPP and CTOPP, we conclude that *Let's Begin with the Letter People* did not have a statistically detectable effect on phonological awareness relative to the control condition.

Language assessments

Data from the two language measures (Peabody Picture Vocabulary Test [PPVT] and Test of Language Development [TOLD] Grammatical Understanding subtest) were analyzed using repeated measures linear spline models. There were no statistically significant differences on these measures at the fall assessment.

There were no statistically detectable differences on any of these measures for the spring pre-kindergarten or spring kindergarten assessments.

Based on analyses of the two language measures, we conclude that *Let's Begin with the Letter People* did not have a statistically detectable effect on language development relative to the control condition.

Behavioral outcomes

We conducted a repeated measures analysis for all three pre-kindergarten social behavioral measures (SSRS Social Skills scale, SSRS Problem Behaviors scale, and PLBS). The covariates were (a) child's age, (b) gender, (c) race/ethnicity, (d) disability status as reported by the parent, and (e) mother's education. There were no statistically detectable differences on these measures for the fall assessment.

For the spring pre-kindergarten assessment, there were no statistically detectable differences between groups on any of these measures.

We analyzed the data from the kindergarten versions of the three behavioral measures (SSRS Social Skills scale, SSRS Problem Behaviors scale, and Learning Behaviors Scale [LBS]) using analysis of covariance (ANCOVA). For the ANCOVA analyses, the covariates included (a) the fall pre-kindergarten score of the pre-kindergarten version of the relevant test, along with (b) child's age, (c) gender, (d) race/ethnicity, (e) disability status as reported by the parent, and (f) mother's education.

There were no statistically detectable differences between groups on any of these measures for the spring kindergarten assessment.

Based on the analyses of the three behavioral measures, we conclude that *Let's Begin with the Letter People* did not have a statistically detectable effect on children's social and learning behaviors relative to the control condition.

Let's Begin with the Letter People—Classroom Outcomes

The unadjusted mean scores for classroom measures are reported in table C-7b in appendix C. Covariate adjusted mean differences and standard errors are reported in table D-7b in appendix D. For all analyses of classroom measures, the following variables were included in the model as covariates: (a) teacher has a BA degree, (b) previous teaching experience, (c) child/adult ratio in classroom, (d) average class size, (e) city size, and (f) geographic site. The classroom-level effect sizes (ES_c) are presented in table 6.5.

Overall classroom environment

We conducted a repeated measures analysis on the ECERS-R. There was no statistically detectable difference between groups on the fall observation.

A statistically significant difference between groups was obtained for the spring pre-kindergarten assessment ($ES_c = .82, p < .05$), such that the *Let's Begin with the Letter People* classrooms received higher global classroom quality ratings relative to the control classrooms.

Based on the analysis of the ECERS-R, we conclude that *Let's Begin with the Letter People* had a positive effect on overall classroom quality relative to the control condition.

Teacher-child relationships

We obtained observations on the Arnett Detachment, Harshness, Permissiveness, and Positive Interactions scales in fall and spring of the pre-kindergarten year, and conducted repeated measures analyses. There was a statistically significant difference on the Arnett Permissiveness scale on the fall observation ($ES_c = .99, p < .01$; follow-up analyses for this finding are included in appendix A). Teachers in the *Let's Begin with the Letter People* classrooms were rated as being more permissive in their interactions with their students relative to teachers in the control classrooms. There were no statistically detectable differences between groups on the other scales for the fall observation.

No statistically detectable differences were obtained for the spring pre-kindergarten assessment on the Arnett Detachment, Permissiveness, and Positive Interactions scales. There was a statistically significant difference

on the Arnett Harshness scale. *Let's Begin with the Letter People* teachers were rated as being less harsh in their interactions with their students relative to teachers in control classrooms ($ES_C = -.95, p < .05$).

Based on the analyses of the four Arnett scales, we conclude that *Let's Begin with the Letter People* did not have a statistically detectable effect on teacher-child relationships relative to the control condition.

Classroom instruction

We obtained observations on classroom instruction in (a) early literacy (TBRS Print and Letter Knowledge and Written Expression scales), (b) phonological awareness (TBRS Phonological Awareness scale), (c) language (TBRS Book Reading and Oral Language scales), and (d) early mathematics (TBRS Math Concepts scale) in spring of pre-kindergarten only. To analyze these data, ANCOVAs were conducted; the covariates were: (a) teacher has a BA degree, (b) previous teaching experience, (c) child/adult ratio in classroom, (d) average class size, (e) city size, and (f) geographic site.

There were no statistically detectable differences on the Book Reading, Written Expression, Phonological Awareness, Oral Language, or Math Concepts scales. There was a statistically significant difference on the Print and Letter Knowledge scale ($ES_C = .99, p < .05$) indicating that the *Let's Begin with the Letter People* teachers provided more instruction on print and letter knowledge relative to teachers in the control classrooms.

Based on the analyses of the TBRS Print and Letter Knowledge and Written Expression scales, we conclude that *Let's Begin with the Letter People* had a positive effect on early literacy instruction relative to the control condition.

Based on the analysis of the TBRS Phonological Awareness scale, we conclude that *Let's Begin with the Letter People* did not have a statistically detectable effect on instruction in phonological awareness relative to the control condition.

Based on the analyses of the TBRS Book Reading and Oral Language scales, we conclude that *Let's Begin with the Letter People* did not have a statistically detectable effect on language instruction relative to the control condition.

Based on the analysis of the TBRS Math Concepts scale, we conclude that *Let's Begin with the Letter People* did not have a statistically detectable effect on early mathematics instruction relative to the control condition.

Summary of Findings for *Let's Begin with the Letter People*

The findings for *Let's Begin with the Letter People* are summarized in table 6.5.

Table 6.5. Effect sizes for *Let's Begin with the Letter People*

Measure	Student-level effect sizes (ES _s)		
	RM analysis Spring Pre-K	RM analysis Spring K	ANCOVA Spring K
Mathematics			
WJ Applied Problems	-.10	-.13	—
CMA-A Mathematics Composite	.15	-.07	—
Shape Composition ¹	.21	-.06	—
Reading			
TERA	.02	-.13	—
WJ Letter Word Identification	.10	-.18	—
WJ Spelling	.17	-.06	—
Phonological awareness			
Pre-CTOPPP/CTOPP	-.13	†	-.13
Language			
PPVT	-.03	.00	—
TOLD	.08	-.12	—
Behavior			
SSRS Social Skills	-.27	†	.24
SSRS Problem Behavior ²	-.06	†	.06
PLBS/LBS	-.44	†	-.10
Measure	Classroom-level effect sizes (ES _c)		
	RM analysis Spring Pre-K	ANCOVA Spring Pre-K	
Global classroom quality			
ECERS-R	.82*	—	
Teacher-child interaction			
Arnett Detachment ³	-.07	—	
Arnett Harshness ³	-.95*	—	
Arnett Permissiveness ³	-.05	—	
Arnett Positive Interactions	.48	—	
Teacher instructional practices ⁴			
TBRS Book Reading	†	.63	
TBRS Oral Language	†	.44	
TBRS Phonological Awareness	†	.66	
TBRS Print and Letter Knowledge	†	.99*	
TBRS Written Expression	†	.60	
TBRS Math Concepts	†	.24	

— Not available.

† Not applicable. Four of the kindergarten student-level measures were not on the same scale as the pre-kindergarten measures. The classroom-level data were only collected during the pre-kindergarten year of the study.

* $p < .05$

¹ Building Blocks, Shape Composition task

² Higher scores on this scale represent more negative child behaviors.

³ Lower scores on this scale represent a more positive classroom environment.

⁴ ANCOVA models for the TBRS measures did not include baseline pretest scores because TBRS data were only collected in spring of the pre-kindergarten year.

NOTE: RM: Repeated Measures

ANCOVA: Analysis of covariance

Significance indications (p -values) in the table refer to the tests of contrasts between intervention and control groups that underlie the effect sizes reported here. Refer to the glossary for abbreviations of the measures.

SOURCE: The Preschool Curriculum Evaluation Research (PCER) Study.

Chapter 7. *Early Literacy and Learning Model (ELLM)*: University of North Florida (Florida-UNF site)

Curriculum

The University of North Florida (Florida-UNF) team implemented *Early Literacy and Learning Model (ELLM)*, a literacy-focused curriculum and support system designed for young children from low-income families. The *ELLM* program components include the following:

- curriculum and literacy building blocks;
- assessment for instructional improvement;
- professional development for literacy coaches and teachers;
- family involvement; and
- collaborative partnerships.

The *ELLM* curriculum and support system is designed to enhance existing classroom curricula by specifically focusing on children's early literacy skills and knowledge.

The *ELLM* curriculum materials include a set of literacy performance standards; monthly literacy packets; targeted instructional strategies; resource guides for teachers; a book lending library; family and teacher tip sheets; and literacy calendars. One hour of daily literacy instruction is required to implement the *ELLM* literacy building blocks. Trained literacy coaches provide instructional support to preschool teachers who use the curriculum.

The *ELLM* program contains a family involvement action plan. Families have access to many resources, including a classroom book-lending library that enables children to take books home daily to share with their parents. Parents receive monthly family tip sheets and calendars with suggestions for literacy activities they can engage in with their children. Parents also have the opportunity to engage in preschool site-based family activities during the school year.

As part of the Florida-UNF complementary study, the *ELLM* program included two evaluation instruments: the *Test of Early Reading Ability-Third Edition, Form A (TERA-3)*³, and the *Alphabet Letter Recognition Inventory (ALRI)* were used as assessment tools. *ELLM* teachers used results from these assessments to identify children's literacy needs and inform classroom literacy instruction. For example, children's fall scores were used to help teachers focus instruction and identify children for targeted instruction in phonological awareness and letter recognition.

Sample

During the 2003-04 academic year, the Florida-UNF research team recruited 28 preschool classrooms from three geographic locations in Florida. The sampled classrooms included Head Start, subsidized, faith-based, and early intervention pre-kindergarten classrooms. All of the classrooms were full-day programs. Twenty-eight classrooms and teachers were recruited to participate in the study. The Florida-UNF research team attended site orientation and/or parent meetings to recruit participants. The teachers and program administrators assisted with the recruitment efforts. Consent forms were sent home to parents and teachers

³ The TERA used in the *ELLM* classrooms was a different version (Form A) than that used for the PCER evaluation study assessment (Form B).

collected signed consent forms from parents. A total of 297 children and 294 parents were recruited for participation in the study. The final sample included 244 children (137 treatment, 107 control) and 243 parents. Data were collected on 243 children and 204 parents at the time of the fall assessment data collection. During the study year, two sites (one control, one intervention) withdrew from the study.

In the follow-up year of the evaluation (the 2004-05 academic year), the sample of schools went from 28 preschools to 119 schools with kindergarten classrooms. The sample of classrooms went from 28 preschool classrooms to 175 kindergarten classrooms. The kindergarten sample included 237 children and 236 parents from original sample of 248 participants. Data were collected on 218 children and 177 parents.

Children and Families

The average age of children was 4.6 years at the time of fall assessment data collection and half (50%) was male. The overall sample was primarily African American (71%) with smaller percentages of White (14%) and Hispanic (8%) children. Table 7.1 provides additional information on the demographic characteristics of the children in the study sample. There were no statistically detectable differences between the treatment and control groups on these child characteristics.

Table 7.1. Child demographic characteristics for *Early Literacy and Learning Model*

Characteristics	Full sample n = 244	Curriculum comparison	
		Control n = 107	Treatment n = 137
Age at baseline (years), mean	4.6	4.6	4.6
Gender (% male)	50.0	48.6	51.1
Race/ethnicity (%)			
White, non-Hispanic	13.8	17.0	11.3
African American, non-Hispanic	71.1	69.1	72.6
Hispanic	7.8	4.3	10.5
Asian or Pacific Islander	‡	‡	‡
Native American	0.0	0.0	0.0
Multiple/other	6.0	7.4	4.8
Child disability status (parent reported, %)	12.7	8.9	15.8

‡ Reporting standards not met. Values suppressed to protect participant confidentiality.

SOURCE: PCER Parent Interview (Fall 2003, Spring 2004, and Spring 2005).

The demographic characteristics of the primary caregivers, who were most often the biological or adoptive mother, are presented in table 7.2. The average age of the primary caregiver was 31 years. Almost half (40%) of the primary caregivers were never married; 37 percent were married at the time of the fall assessment data collection. More than one-third of the primary caregivers reported having had some college (36%) or had graduated from college (6%); 37 percent had a high school diploma or GED; and 22 percent had not finished high school. More than half (54%) of the primary caregivers were employed full-time; 11 percent were employed part-time; and 33 percent were unemployed. At baseline, a higher percentage of parents in the treatment group had completed some post-high school education relative to those assigned to the control group (41% vs. 29%, $p < .01$).

Teachers

There were 28 teachers who participated in the preschool year intervention study. All were female. The majority identified themselves as African American (64%) or White (21%). The preschool teachers had on

Table 7.2. Primary caregiver demographic characteristics for *Early Literacy and Learning Model*

Characteristics	Full sample n = 204	Curriculum comparison	
		Control n = 90	Treatment n = 114
Age at baseline (years), mean	30.9	31.0	30.8
Marital status (%)			
Married	36.8	41.1	33.3
Separated/Divorced	20.6	18.9	21.9
Widowed	2.5	‡	‡
Never Married	40.2	36.7	43.0
Race/ethnicity (%)			
White, non-Hispanic	13.7	16.7	11.4
African American, non-Hispanic	74.0	72.2	75.4
Hispanic	6.4	‡	9.6
Asian or Pacific Islander	2.0	‡	‡
Native American	0.0	0.0	0.0
Multiple/other	3.9	6.7	‡
Educational level (%)			
Did not finish high school	21.6	32.2	13.2
High school diploma or GED	36.8	28.9	43.0
Some college	35.8	28.9	41.2**
College graduate	5.9	10.0	‡
Employment (%)			
Full-time	54.4	45.6	61.4
Part-time	10.8	12.2	9.6
Unemployed	33.3	38.9	28.9
Other	‡	‡	0.0

‡ Reporting standards not met. Values suppressed to protect participant confidentiality.

** $p < .01$

SOURCE: PCER Parent Interview (Fall 2003, Spring 2004, and Spring 2005).

average 11 years of teaching experience, with an average of seven years teaching preschool. Fifty percent of the teachers had a high school diploma or GED and 21 percent had a bachelor's degree. Many of the teachers reported having a state-awarded preschool certificate (52%); a teaching license or certificate (46%); or a Child Development Associate (CDA) credential (46%). Eighteen percent reported having no teacher certification credentials. Table 7.3 provides additional information on the characteristics of the preschool sample of teachers. At baseline, teachers in the treatment group had more years of experience teaching in a preschool setting relative to those assigned to the control group (9 years vs. 4 years, $p < .01$).

Programs/Classrooms

The average preschool class size was 15.5 children. The child-staff ratio was an average of 9.6 children to one teacher or program staff person.

Random Assignment

Randomization was done during the pilot-year study (2002-03). The original random assignment procedure and changes that were made during the evaluation study year are summarized here. A total of 30 classrooms and teachers were included in the pilot-year study sample. Preschool classrooms were randomly assigned to

Table 7.3. Preschool teacher characteristics for Early Literacy and Learning Model

Characteristics	Full sample n = 28	Curriculum comparison	
		Control n = 14	Treatment n = 14
Gender (% female)	100.0	100.0	100.0
Race/ethnicity (%)			
White, non-Hispanic	21.0	29.0	‡
African American, non-Hispanic	64.0	50.0	79.0
Hispanic	‡	‡	‡
Asian or Pacific Islander	0.0	0.0	0.0
Native American	0.0	0.0	0.0
Multiple/other	‡	‡	0.0
Educational level (%)			
High school diploma or GED	50.0	50.0	50.0
Associate's degree	‡	‡	‡
Bachelor's degree	21.0	29.0	‡
Graduate degree	‡	‡	‡
Current teaching license/certificate (%)	46.0	43.0	50.0
Child Development Associate (CDA) (%)	46.0	43.0	50.0
State-awarded preschool certificate (%)	52.0	62.0	43.0
No credential (%)	18.0	‡	‡
Years of teaching experience, overall (mean)	10.7	9.1	12.3
Years of preschool teaching experience (mean)	6.7	4.0	9.4**

‡ Reporting standards not met. Values suppressed to protect participant confidentiality.

** $p < .01$

SOURCE: PCER Preschool Teacher Survey (Fall 2003 and Spring 2004).

treatment or control conditions. The Florida-UNF researchers recruited preschool programs from three distinct geographic locations within the state. The research team first identified elementary school neighborhoods in each geographic location (Counties A, B, and C) with low-performing schools. Using the Florida Department of Education's school grading report card system,⁴ the research team identified grade D and F elementary schools in each of the three counties. It was expected that children from the preschool programs in these low-performing elementary school neighborhoods would transition into these grade D and F elementary schools during the kindergarten year of the study. Preschool programs within the low-performing elementary school neighborhoods were randomly selected for inclusion in the sampling pool of preschool programs.

The sampled preschool classrooms included Head Start, subsidized, faith-based, and early intervention pre-kindergarten programs. Thirty preschool classrooms (10 in County A, 10 in County B, and 10 in County C) were randomly assigned to the treatment or control condition. Only one preschool classroom per preschool was randomly assigned to the treatment or control condition.

⁴ All schools in Florida receive a grade based on the following: (1) percentage of students meeting high standards of the Florida Comprehensive Assessment Test (FCAT)-achievement scores of Level 3 or above; (2) percentage of students making learning gains; and (3) adequate progress of the lowest 25 percent of the students in the school. Each school receives a certain number of points for each of three categories. The points are summed to create a total score. The total score is converted into a letter: grade A (410 points or more), grade B (380 to 409 points), grade C (320 to 379), grade D (280 to 319), and grade F (less than 250). For a grade of A, 95 percent were tested and at least 50 percent of the lowest readers must have made gains in the current school year. For a grade of B or C at least 50 percent of the lowest readers must have made gains in one or two consecutive years. Information reported here is based on the school grade categories in use during the 2003-04 school year. Source: <http://schoolgrades.fldoc.org>.

All of the preschool classrooms in a given elementary school neighborhood were randomly assigned to only one of two conditions (*ELLM* or control). In County A, the research team used a random number software program to randomly assign preschools to conditions. In Counties B and C, all preschools were identified, preschool names were written on strips of paper, and placed in a hat. Preschool classrooms were randomly assigned, one at a time, first to treatment and then to control, until all of the preschools were assigned to one of two conditions.

During the evaluation study year, 14 of the 15 pilot-year treatment classrooms remained in the study. To replace a preschool program that withdrew from the study during the pilot year, a classroom was selected from the site-specific classrooms randomly assigned to implement *ELLM* during the pilot year. Fifteen new control classroom teachers were recruited in the second year of the study to replace those from the pilot year who then received *ELLM* training during the second year as part of their agreement to participate in the pilot study. Preschool programs located within the original elementary school neighborhoods were identified and new control classroom sites were randomly selected from a pool of preschool classrooms in each elementary school neighborhood. Fifteen new control group teachers participated in the study during the intervention year (2003-04). The final evaluation study sample included a total of 28 classrooms (28 of the 30 classrooms remained in the study for the duration of the pre-kindergarten school year) and 299 children.

Contamination

Because all preschool classrooms were assigned to only one of two conditions, there was little risk of contamination across the treatment and control conditions.

Control Condition

A number of curricula were represented in the control classrooms including *Creative Curriculum* (Dodge, Colker, and Heroam 2002), *Beyond Centers and Circletime* (Phelps 2002), *High Reach Learning Pre-K* (High Reach Learning 1997a and 1997b), and *High/Scope* (Hohmann and Weikart 2002).

Data Collection

RTI International (RTI) collected the child, teacher, and school data for the three Florida-UNF sites (Counties A, B, and C) for all three waves of data collection. The Florida-UNF research team was responsible for conducting the parent interviews in the preschool year. In the kindergarten follow-up year, RTI staff completed the parent interviews. The fall assessment data collection window for child assessments ranged from September 8, 2003 to October 30, 2003 (County B); September 4, 2003 to October 22, 2003 (County A); and September 15, 2003 to December 4, 2003 (County C). The average delay from the beginning of the treatment (i.e., start of the school year) to the beginning of the fall assessment window was 27 days in County B, 28 days in County A, and 21 days in County C. The spring pre-kindergarten window was April 2, 2004 to May 7, 2004 (County B); April 13, 2004 to May 6, 2004 (County A); and May 3, 2004 to June 30, 2004 (County C). The kindergarten follow-up window was April 5, 2005 to June 27, 2005 (County B); April 4, 2005 to June 22, 2005 (County A); and April 8, 2005 to June 15, 2005 (County C).

Attrition

Thirty classrooms were randomly assigned to treatment and control conditions. The final sample included 28 classrooms that remained in the study throughout the pre-kindergarten year.

For the child assessment, the fall assessment response rate was 98 percent, the spring 2004 response rate was 92 percent, and the kindergarten follow-up response rate was 92 percent.

Implementation

Eleven of the 14 teachers in the *ELLM* condition were in their second year of implementation of the curriculum at the time of the evaluation.

The *ELLM* literacy curriculum was implemented in combination with the existing comprehensive curricula. Three *ELLM* literacy coaches were trained during a 5-day training session in August 2003. A 2-day follow-up training institute was held in October 2003. Ongoing training of coaches included weekly local seminars at each site location, monthly regional seminars, and monthly regional collaboration team meetings. Teacher training included a 2-day summer training session; weekly classroom visits by *ELLM* literacy coaches; monthly site-specific literacy team meetings; and quarterly teacher get-togethers. Teacher training focused on the *ELLM* curriculum, *ELLM* learning materials, and strategies to help children acquire important emergent literacy skills.

The *ELLM* literacy coaches made weekly literacy visits (1 hour) to intervention classrooms. *ELLM* literacy coaches hosted monthly literacy team meetings at each site location. At the monthly meetings, the literacy coaches distributed monthly materials and resources; demonstrated the use of monthly literacy packets and children's books; shared instructional ideas, and highlighted targeted activities. The teachers also gave the coaches feedback on the effectiveness of their classroom visits and how to better meet the needs of individual teachers.

The Florida-UNF research team collected videotaped data to measure the fidelity of *ELLM* curriculum implementation. Trained videographers videotaped teachers twice (fall 2003 and spring 2004) during the school year. The videotapes were segmented and coded to analyze fidelity of implementation, and were coded to capture the presence or absence of the critical *ELLM* elements in the intervention classrooms. The possible scores on the *ELLM* fidelity-of-use instrument ranged from 0 to 147. A high level of *ELLM* curriculum implementation is defined as 80 percent (118) of possible points on the fidelity-of-use instrument. This level of implementation is aligned with the competent level on the *ELLM* teacher implementation measure. A low level of *ELLM* implementation is reflected by 60 percent (0-88) of possible points on the fidelity-of-use instrument.

Site-Specific Fidelity Ratings

On the site-specific fidelity measure across both assessment times, the intervention classrooms were rated at a Low or Medium level of implementation. No intervention teacher was rated as a high implementer. With one exception, the control classrooms were rated at a Low level of implementation. One control classroom received a Medium level of implementation rating during the spring 2004 observation.

Implementation Fidelity Ratings

Each research team used a global fidelity measure to rate the overall fidelity with which the curricula were implemented in the preschool year of the project. A four-point scale ranging from "Not at All" (0) to "High" (3) was used to rate each treatment classroom. Researchers were asked to use their site-specific implementation and fidelity data to rate each treatment classroom on the global fidelity measure as High, Medium, Low, or Not at All. Researchers were also asked to provide a global rating for the control group curriculum. The *ELLM* curriculum was rated at the Medium implementation fidelity level (2.5). The research team did not provide the RTI evaluation staff with a global fidelity rating (using the four-point scale) for the control group classrooms at their research site.

Impact Analysis Results

We begin with the analyses of the child-level measures (i.e., the mathematics, reading, phonological awareness, and language assessments) and then present the analyses of the classroom observation data. Our discussion of the results focuses on the combined analysis of the three sites.

Early Literacy and Learning Model—Child Outcomes

The unadjusted mean scores for child-level measures are reported in table C-8a in appendix C. Covariate adjusted mean differences and standard errors are reported in table D-8a in appendix D. For all analyses of child-level measures, the following covariates were included: (a) child's age, (b) gender, (c) race/ethnicity, (d) disability status as reported by parent, and (e) mother's education. The student-level effect sizes (ES_s) are presented in table 7.4.

Mathematics assessments

We used repeated measures linear spline models to analyze the data from all three mathematics measures (Woodcock Johnson [WJ] Applied Problems, Child Math Assessment-Abbreviated [CMA-A] Composite Score, and Shape Composition). There were no statistically detectable differences at the fall assessment.

There were no statistically detectable differences between groups on any of these measures for the spring pre-kindergarten or kindergarten assessments.

Based on the analyses for the three mathematics measures, we conclude that *ELLM* did not have a statistically detectable effect on mathematics relative to the control condition.

Reading assessments

Data from the three reading measures (Test of Early Reading Ability [TERA], WJ Letter Word Identification, and WJ Spelling) were analyzed using repeated measures linear spline models. There were no statistically detectable differences on these measures for the fall assessment.

There were no statistically detectable differences between groups on any of these measures for the spring pre-kindergarten or kindergarten assessments.

Based on the analyses for the three reading measures, we conclude that *ELLM* did not have a statistically detectable effect on reading relative to the control condition.

Phonological awareness

The phonological awareness measures were the Preschool Comprehensive Test of Phonological and Print Processing (Pre-CTOPPP), Elision subtest, and the Comprehensive Test of Phonological Processing (CTOPP), Kindergarten, Elision subtest. We conducted a repeated measures analysis on the Pre-CTOPPP fall and spring pre-kindergarten data. There was no statistically detectable difference on the Pre-CTOPPP for the fall assessment.

There was no statistically detectable difference on the Pre-CTOPPP for the spring pre-kindergarten assessment.

We analyzed the kindergarten CTOPP data using analysis of covariance (ANCOVA). For the ANCOVA analysis, the covariates were the (a) Pre-CTOPPP fall assessment score, (b) child's gender, (c) age, (d) race/ethnicity, (e) disability status as reported by parent, and (f) mother's education. There was no statistically significant difference between groups on the CTOPP for the spring kindergarten assessment.

Based on the analyses of the Pre-CTOPPP and CTOPP, we conclude that *ELLM* did not have a statistically detectable effect on phonological awareness relative to the control condition.

Language assessments

Data from the two language measures (Peabody Picture Vocabulary Test [PPVT] and Test of Language Development [TOLD] Grammatical Understanding subtest) were analyzed using repeated measures linear spline models. There were no statistically detectable differences on either measure for the fall assessment.

For the spring pre-kindergarten assessment, there were no statistically significant mean differences between groups on either measure. However, for the spring kindergarten assessment, there were statistically reliable differences on the PPVT ($ES_s = .34, p < .05$) and the TOLD Grammatical Understanding ($ES_s = .44, p < .05$).

Based on the analyses of the two language measures, we conclude that *ELLM* did not have a statistically detectable effect on language development relative to the control condition in the pre-kindergarten year. However, results indicate there was a delayed effect of *ELLM* on language development relative to the control condition at the end of the kindergarten year.

Behavioral outcomes

We conducted a repeated measures analysis for all three pre-kindergarten social behavioral measures (Social Skills Rating System [SSRS] Social Skills scale, SSRS Problem Behaviors scale, and Preschool Learning Behaviors Scale [PLBS]). The covariates were (a) child's age, (b) gender, (c) race/ethnicity, (d) disability status as reported by the parent, and (e) mother's education. There were no statistically detectable differences on the behavioral measures for the fall assessment.

For the spring pre-kindergarten assessment, there were no statistically detectable differences on any of these measures.

We analyzed the data from the kindergarten versions of the three behavioral measures (SSRS Social Skills scale, SSRS Problem Behaviors scale, and Learning Behaviors Scale [LBS]) using analysis of covariance (ANCOVA). For the ANCOVA analyses, the covariates included (a) the fall pre-kindergarten score of the pre-kindergarten version of the relevant test, along with (b) child's age, (c) gender, (d) race/ethnicity, (e) disability status as reported by the parent, and (f) mother's education.

There were no statistically detectable differences between groups on any of these measures for the spring kindergarten assessment.

Based on the analyses of the three behavioral measures, we conclude that *ELLM* did not have a statistically detectable effect on children's social and learning behaviors relative to the control condition.

Early Literacy and Learning Model—Classroom Outcomes

The unadjusted mean scores for classroom measures are reported in table C-8b in appendix C. Covariate adjusted mean differences and standard errors are reported in table D-8b in appendix D. For all analyses of classroom measures, the following variables were included in the model as covariates: (a) teacher has a BA degree, (b) previous teaching experience, (c) child/adult ratio in classroom, (d) average class size, (e) city size, and (f) geographic site. The classroom-level effect sizes (ES_c) are presented in table 7.4.

Overall classroom environment

We conducted a repeated measures analysis on the Early Childhood Environment Rating Scale-Revised (ECERS-R). There was no statistically detectable difference between groups for the fall observation. No statistically detectable difference between groups was obtained for the spring pre-kindergarten observation.

Based on the analysis of the ECERS-R, we conclude that *ELLM* did not have a statistically detectable effect on overall classroom quality relative to the control condition.

Teacher-child relationships

We obtained observations on the Arnett Detachment, Harshness, Permissiveness, and Positive Interactions scales in fall and spring of the pre-kindergarten year, and conducted repeated measures analyses. There were no statistically detectable differences on these measures in fall of the pre-kindergarten year.

There were no statistically detectable differences between groups on the Arnett Detachment, Harshness, Permissiveness, or Positive Interaction scales for the spring pre-kindergarten observation.

Based on the analyses of the four Arnett scales, we conclude that *ELLM* did not have a statistically detectable effect on teacher-child relationships relative to the control condition.

Classroom instruction

We obtained observations on classroom instruction in (a) early literacy (Teacher Behavior Rating Scale [TBRS] Print and Letter Knowledge and Written Expression scales), (b) phonological awareness (TBRS Phonological Awareness scale), (c) language (TBRS Book Reading and Oral Language scales), and (d) early mathematics (TBRS Math Concepts scale) in spring of pre-kindergarten only. To analyze these data, ANCOVAs were conducted; the covariates were: (a) teacher has a BA degree, (b) previous teaching experience, (c) child/adult ratio in classroom, (d) average class size, (e) city size, and (f) geographic site.

There were no statistically reliable differences between groups on any of the TBRS scales.

Based on the analyses of the TBRS Print and Letter Knowledge and Written Expression scales, we conclude that *ELLM* did not have a statistically detectable effect on early literacy instruction relative to the control condition.

Based on the analysis of the TBRS Phonological Awareness scale, we conclude that *ELLM* did not have a statistically detectable effect on instruction in phonological awareness relative to the control condition.

Based on the analysis of the TBRS Book Reading and Oral Language scales, we conclude that *ELLM* did not have a statistically detectable effect on language instruction relative to the control condition.

Based on the analysis of the TBRS Math Concepts scale, we conclude that *ELLM* did not have a statistically detectable effect on early mathematics instruction relative to the control condition.

Summary of Findings for *Early Literacy and Learning Model*

The findings for *ELLM* are summarized in table 7.4.

Table 7.4. Effect sizes for Early Literacy and Learning Model

Measure	Student-level effect sizes (ES _s)		
	RM analysis Spring Pre-K	RM analysis Spring K	ANCOVA Spring K
Mathematics			
WJ Applied Problems	.10	.26	—
CMA-A Mathematics Composite	.01	-.05	—
Shape Composition ¹	-.14	.03	—
Reading			
TERA	.15	.30	—
WJ Letter Word Identification	-.05	.00	—
WJ Spelling	.11	.04	—
Phonological awareness			
Pre-CTOPPP/CTOPP	.18	†	.08
Language			
PPVT	.17	.34*	—
TOLD	.15	.44**	—
Behavior			
SSRS Social Skills	-.06	†	.27
SSRS Problem Behavior ²	-.24	†	.23
PLBS/LBS	.14	†	.04
Measure	Classroom-level effect sizes (ES _c)		
	RM analysis Spring Pre-K	ANCOVA Spring Pre-K	
Global classroom quality			
ECERS-R	-.48	—	
Teacher-child interaction			
Arnett Detachment ³	-.41	—	
Arnett Harshness ³	-.40	—	
Arnett Permissiveness ³	-.24	—	
Arnett Positive Interactions	.29	—	
Teacher instructional practices ⁴			
TBRS Book Reading	†	.32	
TBRS Oral Language	†	.14	
TBRS Phonological Awareness	†	.53	
TBRS Print and Letter Knowledge	†	.41	
TBRS Written Expression	†	-.22	
TBRS Math Concepts	†	-.92	

— Not available.

† Not applicable. Four of the kindergarten student-level measures were not on the same scale as the pre-kindergarten measures. The classroom-level data were only collected during the pre-kindergarten year of the study.

* $p < .05$; ** $p < .01$ ¹ Building Blocks, Shape Composition task² Higher scores on this scale represent more negative child behaviors.³ Lower scores on this scale represent a more positive classroom environment.⁴ ANCOVA models for the TBRS measures did not include baseline pretest scores because TBRS data were only collected in spring of the pre-kindergarten year.

NOTE: RM: Repeated Measures

ANCOVA: Analysis of covariance

Significance indications (p -values) in the table refer to the tests of contrasts between intervention and control groups that underlie the effect sizes reported here. Refer to the glossary for abbreviations of the measures.

SOURCE: The Preschool Curriculum Evaluation Research (PCER) Study.

Chapter 8. *Language-Focused Curriculum*: University of Virginia (Virginia site)

Curriculum

The University of Virginia (Virginia) research team evaluated the *Language-Focused Curriculum (LFC)*. The *LFC* was developed through a 1985 Model Demonstration Project funded by the U.S. Department of Education to the University of Kansas to design a Language Acquisition Preschool. The *LFC* was designed for use with 3- to 5-year-old children with language limitations, including children with language impairment; children from disadvantaged backgrounds; and English-language learners.

The curriculum components include the following:

- thematic organization of content by day, week, and month;
- use of daily dramatic play to teach and use new linguistic concepts;
- use of both teacher-led and child-led activities to organize daily experiences;
- explicit attention to oral language goals across the day; and
- teacher use of the eight key “language stimulation techniques” when interacting with children in the classroom.

The *LFC* emphasizes the daily inclusion of high-quality teacher-child conversations within teacher-led and child-led interactions.

Sample

The Virginia research team recruited 14 teachers and preschool classrooms to participate in the study. A combination of Head Start and public pre-kindergarten classrooms was recruited. All of the programs were full-day programs. Teachers received incentives for participating in the study. Teachers and school administrators assisted with the recruitment of parents and children. The parent and child recruitment process occurred during the first few weeks of the school year. An incentive (storybooks) was offered to children as part of the parental consenting process. A total sample of 205 children and parents were recruited for the study. The average parental consent rate was 94 percent (95% for the treatment group, 93% for the control group). The final sample included 195 children (97 treatment, 98 control) and parents. Data were collected on 182 children and 181 parents at the time of the fall baseline data collection.

In the follow-up year of the evaluation, the sample of schools went from five in pre-kindergarten to 21 schools in kindergarten. The sample of classrooms went from 14 preschool to 54 kindergarten classrooms. The kindergarten sample included 189 of the original sample of 195 children. Data were collected on 189 children and 174 parents.

Children and Families

The children were 4.6 years of age at the time of baseline data collection and slightly more than half (53%) were male. The majority of the sample of preschoolers were White (71%) or African American (21%). Table 8.1 provides additional information on the demographic characteristics of the children in the Virginia study sample. There were no statistically detectable differences between the treatment and control groups on these child characteristics.

Table 8.1. Child demographic characteristics for *Language-Focused Curriculum*

Characteristics	Full sample n = 195	Curriculum comparison	
		Control n = 98	Treatment n = 97
Age at baseline (years), mean	4.6	4.6	4.6
Gender (% male)	52.7	52.7	52.8
Race/ethnicity (%)			
White, non-Hispanic	70.8	67.4	74.4
African American, non-Hispanic	20.8	25.0	16.3
Hispanic	4.5	5.4	‡
Asian or Pacific Islander	0.0	0.0	0.0
Native American	‡	0.0	‡
Multiple/other	2.8	‡	‡
Child disability status (parent reported, %)	17.7	16.1	19.3

‡ Reporting standards not met. Values suppressed to protect participant confidentiality.

SOURCE: PCER Parent Interview (Fall 2003, Spring 2004, and Spring 2005).

The demographic characteristics of the primary caregivers, who were most often the biological or adoptive mother, are presented in table 8.2. The average age of the primary caregiver was 30 years. More than half (54%) of the primary caregivers were married, and 22 percent were never married. Almost half (45%) reported having a high school diploma or GED; 20 percent had not finished high school; 29 percent had some college education; and 7 percent had a BA. Less than half (46%) of the primary caregivers were employed full-time, 39 percent were unemployed, and 14 percent were employed part-time. There were no statistically detectable differences between the treatment and control groups on the primary caregiver characteristics.

Teachers

There were 14 teachers who participated in the preschool-year intervention study. All of the preschool teachers were female, and all were White. On average, the preschool teachers had 11 years of teaching experience, with an average of 8 years of experience teaching preschool. The majority of teachers had a bachelor's (71%) degree. The majority of teachers reported having a state teacher certificate (71%). Table 8.3 provides additional information on the characteristics of the preschool sample of teachers. There were no statistically detectable differences between the treatment and control groups on the teacher characteristics.

Programs/Classrooms

The average preschool class size was 13 children. The child-staff ratio was on average 6.3 children to one teacher or program staff person.

Random Assignment

The research team identified and recruited a convenience sample of preschools from two counties in Virginia (one rural county and one suburban county). Along with the Virginia researchers, Mathematica Policy Research, Inc. (MPR) determined the unit of random assignment for this research site. The MPR research staff randomly assigned individual classrooms to conditions after it was determined that the experimental curriculum could be introduced in one classroom without affecting neighboring classrooms in the same school and, second, that preschool staff were willing to use different curricula in the same setting. Individual classrooms within schools were randomly assigned to treatment and control conditions. To increase the

Table 8.2. Primary caregiver demographic characteristics for *Language-Focused Curriculum*

Characteristics	Full sample n = 179	Curriculum comparison	
		Control n = 93	Treatment n = 86
Age at baseline (years), mean	29.8	30.5	29.2
Marital status (%)			
Married	54.2	50.5	58.1
Separated/Divorced	22.3	22.6	22.1
Widowed	‡	‡	0.0
Never Married	22.3	24.7	19.8
Race/ethnicity (%)			
White, non-Hispanic	74.9	76.3	73.3
African American, non-Hispanic	18.4	20.4	16.3
Hispanic	5.0	‡	7.0
Asian or Pacific Islander	0.0	0.0	0.0
Native American	‡	0.0	‡
Multiple/other	‡	0.0	‡
Educational level (%)			
Did not finish high school	19.6	22.6	16.3
High school diploma or GED	44.7	40.9	48.8
Some college	29.1	31.2	26.7
College graduate	6.7	5.4	8.1
Employment (%)			
Full-time	46.4	47.3	45.3
Part-time	14.0	11.8	16.3
Unemployed	39.1	39.8	38.4
Other	‡	‡	0.0

‡ Reporting standards not met. Values suppressed to protect participant confidentiality.

SOURCE: PCER Parent Interview (Fall 2003, Spring 2004, and Spring 2005).

precision with which to estimate impacts, MPR grouped classrooms into blocks of two and randomly assigned half the classrooms in each block to the treatment group and half to the control group. The MPR research staff formed blocks by matching classrooms on easily measured characteristics such as teachers' experience, school location, or score on a state report card system and, in doing so, increased the probability that those characteristics would be evenly distributed between the overall treatment and control groups. MPR staff used a random number function (RAND function in MS Excel) to generate random numbers. They sorted the classrooms by block and assigned a random number to each classroom. The classrooms were then randomly assigned to treatment and control conditions. The staff assigned the highest-ranking classroom within the block to the treatment condition, the next highest to the control condition, alternating assignment to treatment and control conditions until all classrooms were randomly assigned to one of two conditions. A total of 14 classrooms (7 treatment and 7 control) were randomly assigned to conditions.

Table 8.3. Preschool teacher characteristics for *Language-Focused Curriculum*

Characteristics	Full sample n = 14	Curriculum comparison	
		Control n = 7	Treatment n = 7
Gender (% female)	100.0	100.0	100.0
Race/ethnicity (%)			
White, non-Hispanic	100.0	100.0	100.0
African American, non-Hispanic	0.0	0.0	0.0
Hispanic	0.0	0.0	0.0
Asian or Pacific Islander	0.0	0.0	0.0
Native American	0.0	0.0	0.0
Multiple/other	0.0	0.0	0.0
Educational level (%)			
High school diploma or GED	‡	‡	‡
Associate's degree	‡	‡	0.0
Bachelor's degree	71.0	57.0	86.0
Graduate degree	‡	‡	0.0
Current teaching license/certificate (%)	71.0	71.0	71.0
Child Development Associate (CDA) (%)	‡	‡	‡
State-awarded preschool certificate (%)	29.0	‡	‡
No credential (%)	‡	‡	‡
Years of teaching experience, overall (mean)	11.4	11.4	11.3
Years of preschool teaching experience (mean)	8.0	7.4	8.6

‡ Reporting standards not met. Values suppressed to protect participant confidentiality.

SOURCE: PCER Preschool Teacher Survey (Fall 2003 and Spring 2004).

Contamination

In each of the five participating schools, there were both treatment and control classrooms. To reduce the possibility of contamination across conditions, the researchers monitored the classrooms to ensure that treatment group teachers were not sharing materials and instructional practices with the control group teachers.

Control Condition

In the control condition, the teachers reported using *High/Scope* curriculum materials, but the extent of *High/Scope* curriculum implementation in the control classrooms was not formally assessed.

Data Collection

MPR collected the child, parent, teacher, and school data for the Virginia site for all three waves of data collection. The fall assessment data collection window for child assessments ranged from September 29, 2003 to November 11, 2003. The average delay from the beginning of the treatment (i.e., start of the school year) to the beginning of the fall assessment window was 28 days. The spring pre-kindergarten window was April 1, 2004 to June 18, 2004, and the kindergarten follow-up window was March 29, 2005 to June 8, 2005.

Attrition

Fourteen classrooms were randomly assigned to treatment or control condition. All 14 classrooms remained in the study from the beginning of the pre-kindergarten year through the spring of the pre-kindergarten year.

For the child assessment, the baseline (fall 2003) response rate was 93 percent, the spring 2004 response rate was 96 percent, and the kindergarten follow-up response rate was 97 percent.

Implementation

Seven classrooms were assigned to implement *LFC* and seven classrooms maintained the prevailing curriculum (*High/Scope*). Five of the seven teachers and seven teaching assistants completed a 3-day training workshop on *LFC* implementation in August 2003. The workshop content included background information on language development. A one-on-one make-up training session was provided to the remaining two teachers who could not attend the initial workshop. Additional follow-up training sessions were held in November 2003, and January/February 2004. In November 2003 treatment group teachers attended an informal on-site 2-hour workshop to discuss teachers' concerns with the implementation of the *LFC*, review the feedback from the first round of classroom observations, and review language stimulation techniques and appropriate use. All of the teachers attended a 3-hour workshop in January/February 2004. The workshop topic was *Being a Conversational Partner*, which focused on language stimulation in the *LFC*, with periodic follow-up training sessions for further discussion and description of implementation activities. All teachers maintained professional development logs throughout the school year to evaluate the extent of professional development experienced by treatment and control group teachers. Site-specific curriculum fidelity observations were conducted in treatment and control classrooms in the fall and spring of the preschool year.

Implementation Fidelity Ratings

Each research team used a global fidelity measure to rate the overall fidelity with which the curricula were implemented in the preschool year of the project. A four-point scale ranging from “Not at All” (0) to “High” (3) was used to rate each treatment classroom. Researchers were asked to use their site-specific implementation and fidelity data to rate each treatment classroom on the global fidelity measure as High, Medium, Low, or Not at All. Researchers were also asked to provide a global rating for the control group curriculum. Both the *LFC* and the control group curriculum were rated at the Medium (2.0) level on the global implementation fidelity measure.

Impact Analysis Results

We begin with the analyses of the child outcomes (i.e., mathematics, reading, phonological awareness, and language assessments) and then present the analyses of the classroom observation data.

Language-Focused Curriculum—Child Outcomes

The unadjusted mean scores for child-level measures are reported in table C-9a in appendix C. Covariate adjusted mean differences and standard errors are reported in table D-9a in appendix D. For all analyses of child-level measures, the following covariates were included: (a) child's age, (b) gender, (c) race/ethnicity, (d) disability status as reported by parent, and (e) mother's education. The student-level effect sizes (ESs) are presented in table 8.4.

Mathematics assessments

We used repeated measures linear spline models to analyze the data from all three mathematics measures (Woodcock Johnson [WJ] Applied Problems, Child Math Assessment-Abbreviated [CMA-A] Composite Score, and Shape Composition). There were no statistically detectable differences on these measures for the fall assessment.

There were no statistically detectable differences between groups on the spring pre-kindergarten or kindergarten assessments.

Based on the analyses for the three mathematics measures, we conclude that the *LFC* did not have a statistically detectable effect on mathematics relative to the control condition.

Reading assessments

Data from the three reading measures (Test of Early Reading Ability [TERA], WJ Letter Word Identification, and WJ Spelling) were analyzed using repeated measures linear spline models. There were no statistically detectable differences on these measures for the fall assessment.

There were no statistically detectable differences between groups on any of these measures for the spring pre-kindergarten or spring kindergarten assessments.

Based on the analyses for the three reading measures, we conclude that the *LFC* did not have a statistically detectable effect on reading relative to the control condition.

Phonological awareness

The phonological awareness measures were the Preschool Comprehensive Test of Phonological and Print Processing (Pre-CTOPPP), Elision subtest, and the Comprehensive Test of Phonological Processing (CTOPP), Kindergarten, Elision subtest. We conducted a repeated measures analysis on the Pre-CTOPPP fall and spring pre-kindergarten data. There was no statistically detectable difference on the Pre-CTOPPP for the fall assessment.

There was no statistically detectable difference between groups on the Pre-CTOPPP for the spring pre-kindergarten assessment.

We analyzed the kindergarten CTOPP data using analysis of covariance (ANCOVA). For the ANCOVA analysis, the covariates were the (a) Pre-CTOPPP fall assessment score, (b) child's gender, (c) age, (d) race/ethnicity, (e) disability status as reported by parent, and (f) mother's education. There was no statistically detectable difference between groups on the CTOPP for the spring kindergarten assessment.

Based on the analyses of the Pre-CTOPPP and CTOPP, we conclude that the *LFC* did not have a statistically detectable effect on phonological awareness relative to the control condition.

Language assessments

Data from the two language measures (Peabody Picture Vocabulary Test [PPVT] and Test of Language Development [TOLD] Grammatical Understanding subtest) were analyzed using repeated measures linear spline models. There were no statistically detectable differences on these measures for the fall assessment.

There were no statistically detectable differences between groups on any of these measures for the spring pre-kindergarten and spring kindergarten assessments.

Based on the analyses for the two language measures, we conclude that the *LFC* did not have a statistically detectable effect on language development relative to the control condition.

Behavioral outcomes

We conducted a repeated measures analysis for all three pre-kindergarten social behavioral measures (Social Skills Rating System [SSRS] Social Skills scale, SSRS Problem Behaviors scale, and Preschool Learning Behaviors Scale [PLBS]). The covariates were (a) child's age, (b) gender, (c) race/ethnicity, (d) disability status as reported by the parent, and (e) mother's education. There were no statistically detectable differences on these measures for the fall assessment.

For the spring pre-kindergarten assessment, there were no statistically detectable differences on any of these measures.

We analyzed the data from the kindergarten versions of the three behavioral measures (SSRS Social Skills scale, SSRS Problem Behaviors scale, and Learning Behaviors Scale [LBS]) using analysis of covariance (ANCOVA). For the ANCOVA analyses, the covariates included (a) the fall pre-kindergarten score of the

pre-kindergarten version of the relevant test, along with (b) child's age, (c) gender, (d) race/ethnicity, (e) disability status as reported by the parent, and (f) mother's education.

There were no statistically detectable differences between groups on any of these measures for the spring kindergarten assessment.

Based on the analyses of the three behavioral measures, we conclude that the *LFC* did not have a statistically detectable effect on children's social and learning behaviors relative to the control condition.

Language-Focused Curriculum—Classroom Outcomes

The unadjusted mean scores for classroom measures are reported in table C-9b in appendix C. Covariate adjusted mean differences and standard errors are reported in table D-9b in appendix D. For all analyses of classroom measures, the following variables were included in the model as covariates: (a) teacher has a BA degree, (b) previous teaching experience, (c) child/adult ratio in classroom, (d) average class size, (e) city size, and (f) geographic site.

The results from the analysis of the overall classroom environment Early Childhood Environment Rating Scale-Revised (ECERS-R) and teacher-child relationships (Arnett measure) for the Virginia site are not included in this report because of data integrity concerns. During the baseline data collection, one observer completed the observational ratings in 8 of the 12 classrooms at this research site. It was later determined that the ECERS-R and Arnett ratings from these eight classrooms were inflated. Due to concerns regarding the integrity of the data from these eight classrooms, the decision was made to exclude the classroom quality and teacher-child relationships data for this site from the report.

Classroom instruction

We obtained observations on classroom instruction in (a) early literacy (Teacher Behavior Rating Scale [TBRS] Print and Letter Knowledge and Written Expression scales), (b) phonological awareness (TBRS Phonological Awareness scale), (c) language (TBRS Book Reading and Oral Language scales), and (d) early mathematics (TBRS Math Concepts scale) in spring of pre-kindergarten only. To analyze these data, ANCOVAs were conducted; the covariates were (a) teacher has a BA degree, (b) previous teaching experience, (c) child/adult ratio in classroom, (d) average class size, (e) city size, and (f) geographic site. The classroom-level effect sizes (ES_C) are presented in table 8.4.

There were no statistically detectable differences between groups on the TBRS scales.

Based on the analyses of the TBRS Print and Letter Knowledge and Written Expression scales, we conclude that the *LFC* did not have statistically detectable effect on early literacy instruction relative to the control condition.

Based on the analysis of the TBRS Phonological Awareness scale, we conclude that the *LFC* did not have a statistically detectable effect on instruction in phonological awareness relative to the control condition.

Based on the analysis of the TBRS Book Reading and Oral Language scales, we conclude that the *LFC* did not have a statistically detectable effect on language instruction relative to the control condition.

Based on the analysis of the TBRS Math Concepts scale, we conclude that the *LFC* did not have a statistically detectable effect on early mathematics instruction relative to the control condition.

Summary of Findings for Language-Focused Curriculum

The findings for *LFC* are summarized in table 8.4.

Table 8.4. Effect sizes for Language-Focused Curriculum

Measure	Student-level effect sizes (ES _s)		
	RM analysis Spring Pre-K	RM analysis Spring K	ANCOVA Spring K
Mathematics			
WJ Applied Problems	.20	.11	—
CMA-A Mathematics Composite	.08	.00	—
Shape Composition ¹	.08	.06	—
Reading			
TERA	.16	.05	—
WJ Letter Word Identification	.11	.02	—
WJ Spelling	.25	.11	—
Phonological awareness			
Pre-CTOPPP/CTOPP	.20	†	.03
Language			
PPVT	.02	-.09	—
TOLD	.01	-.07	—
Behavior			
SSRS Social Skills	-.42	†	-.07
SSRS Problem Behavior ²	.37	†	-.05
PLBS/LBS	-.27	†	.10
Measure	Classroom-level effect sizes (ES _c)		
	RM analysis Spring Pre-K	ANCOVA Spring Pre-K	
Global classroom quality			
ECERS-R	—	—	
Teacher-child interaction			
Arnett Detachment ³	—	—	
Arnett Harshness ³	—	—	
Arnett Permissiveness ³	—	—	
Arnett Positive Interactions	—	—	
Teacher instructional practices ⁴			
TBRS Book Reading	†	-.79	
TBRS Oral Language	†	.87	
TBRS Phonological Awareness	†	.92	
TBRS Print and Letter Knowledge	†	.33	
TBRS Written Expression	†	.99	
TBRS Math Concepts	†	.20	

— Not available. Data were collected but not reported.

† Not applicable. Four of the kindergarten student-level measures were not on the same scale as the pre-kindergarten measures. The classroom-level data were only collected during the pre-kindergarten year of the study.

¹ Building Blocks, Shape Composition task

² Higher scores on this scale represent more negative child behaviors.

³ Lower scores on this scale represent a more positive classroom environment.

⁴ ANCOVA models for the TBRS measures did not include baseline pretest scores because TBRS data were only collected in spring of the pre-kindergarten year.

NOTE: RM: Repeated Measures

ANCOVA: Analysis of covariance

Refer to the glossary for abbreviations of the measures.

SOURCE: The Preschool Curriculum Evaluation Research (PCER) Study.

Chapter 9. Literacy Express and DLM Early Childhood Express supplemented with Open Court Reading Pre-K: Florida State University (Florida-FSU site)

Curriculum

The Florida State University (Florida-FSU) research team chose to evaluate two curricula: *Literacy Express* and *DLM Early Childhood Express supplemented with Open Court Reading Pre-K*.

Literacy Express

The Florida-FSU research team implemented the *Literacy Express* curriculum. *Literacy Express* is a preschool literacy-focused curriculum that is designed to promote children's emergent literacy skills. The curriculum is structured around thematic units. The units, and the games and activities within each unit, are sequenced in order of complexity. Each thematic unit of the curriculum includes selected children's books that address theme-relevant vocabulary for small- and large-group reading activities. In addition, each thematic unit includes small-group activities that provide children with the opportunity to attend to and practice the skills needed to develop oral language, phonological sensitivity, and print awareness, and to receive individual feedback needed to master each developmental level. Small-group activities are conducted 3-4 times a week. The curriculum provides guidance to teachers on grouping children who are progressing at similar rates. The large-group and extension activities provide opportunities for children to use new skills in novel and varied contexts.

DLM Early Childhood Express supplemented with Open Court Reading Pre-K

The Florida-FSU research team implemented the Open Court literacy-focused curriculum in conjunction with *DLM Early Childhood Express* comprehensive curriculum. The *Open Court Reading Pre-K* curriculum is a literacy-focused curriculum. The curriculum content is presented in eight thematic units that address children's identity, families, friends, social interactions, transportation, the physical senses, nature, and transitions. Phonological, phonemic, and print-awareness activities are incorporated into each lesson. Comprehension activities are also included in each lesson to help promote children's understanding of literature. Each day, teachers read literature selections that focus on the topic that is in a thematic unit. The curriculum includes a home component to encourage home/school connections by providing parents with suggestions for activities that they can engage in at home with their children.

The *DLM Early Childhood Express* Program is a comprehensive curriculum. The *DLM Early Childhood Express* curriculum is designed to promote children's social, emotional, intellectual, aesthetic, and physical development through the use of hands-on learning experiences. The curriculum has 36 weekly themes that address the following content areas: literacy, mathematics, science, social studies, fine arts, health/safety, personal/social development, physical movement, and technology. Each thematic unit includes more than 200 age-appropriate, hands-on learning activities that are designed to promote children's social, emotional, intellectual, aesthetic, and physical development.

By integrating the research-based instruction from *Open Court Reading Pre-K* with the comprehensive instructional framework of *DLM Early Childhood Express*, children received instruction that is intended to provide them with a strong foundation in oral language and print awareness as well as research-based instruction in phonics and early decoding and comprehension skills.

Sample

The description of the recruitment process applies to both curricula. The research team recruited public pre-kindergarten programs for participation in the study. Principals from elementary schools were provided information regarding the proposed project and invited to participate. Two teachers from each of the 16 participating schools were recruited to participate in the study. All of the programs were full-day programs. No incentives were offered to teachers. The final study sample included 30 teachers and classrooms across three conditions (9 control, 10 *Literacy Express*, and 11 *DLM Early Childhood Express supplemented with Open Court Reading Pre-K*). Teachers assisted with the recruitment of parents and children to participate in the study. The parental consent process began at the beginning of the school year and continued into the first few weeks of school. The average parental consent rate was 94 percent (95% for the treatment group; 93 percent for the control group). A total of 297 children (99 in the *Literacy Express* treatment group; 101 in the *DLM Early Childhood Express supplemented with Open Court Reading Pre-K* treatment group; and 97 in the control group) and parents were recruited. Data were collected on a total of 282 children and 270 parents at the time of the fall baseline data collection.

In the follow-up year of the evaluation, the sample of schools went from 17 in pre-kindergarten to 46 schools in kindergarten. The sample of classrooms went from 30 preschool to 145 kindergarten classrooms. Data were collected on 237 children and 223 parents from the original sample.

Children and Families

The children were 4.6 years of age at the time of baseline data collection and slightly more than half (54%) were male. The majority of the sample of preschoolers was African American (59%) or White (30%). Table 9.1 provides additional information on the demographic characteristics of the children in the study sample. There were no statistically detectable differences between the treatment and control groups on these child characteristics.

The demographic characteristics of the primary caregivers, who were most often the biological or adoptive mother, are presented in table 9.2. The average age of the primary caregiver was 31 years. About one-third (36%) were married, and 43 percent were never married. Approximately one-third (34%) of the primary caregivers reported having a high school diploma or GED; 13 percent had not finished high school; 38 percent had some college education; and 15 percent had a bachelor's degree or higher. More than half (63%) of the primary caregivers were employed full-time, 12 percent were employed part-time, and 23 percent were unemployed. There were no statistically detectable differences between the treatment and control groups on the primary caregiver characteristics.

Teachers

There were 30 teachers who participated in the preschool-year intervention study. Most (97%) were female, and most were White (83%) or African American (13%). On average, the preschool teachers had 16 years of teaching experience, with an average of 9 years of experience teaching preschool. The majority of teachers had a bachelor's (53%) or graduate (27%) degree. An additional 13 percent had a high school diploma or GED. The majority of teachers reported having a current teaching license or certificate (80%). Some teachers also had a state-awarded preschool certificate (40%), or a Child Development Associate (CDA) credential (23%). Table 9.3 provides additional information on the characteristics of the preschool sample of teachers. There were no statistically detectable differences between the treatment and control groups on the teacher characteristics.

Programs/Classrooms

The average preschool class size was 14 children. The child-staff ratio was on average 5.7 children to one teacher or program staff person in both locations.

Table 9.1. Child demographic characteristics for *Literacy Express* and *DLM Early Childhood Express* supplemented with *Open Court Reading Pre-K*

Characteristics	Full sample n = 297	Curriculum comparison		
		Control n = 97	Treatment 1 ¹ n = 99	Treatment 2 ² n = 101
Age at baseline (years), mean	4.6	4.6	4.6	4.6
Gender (% male)	54.3	59.3	52.7	51.0
Race/ethnicity (%)				
White, non-Hispanic	29.6	23.5	40.2	25.5
African American, non-Hispanic	58.9	70.6	50.6	56.1
Hispanic	5.6	‡	5.7	8.2
Asian or Pacific Islander	‡	0.0	0.0	‡
Native American	‡	‡	‡	‡
Multiple/other	4.8	‡	‡	9.2
Child disability status (parent reported, %)	35.9	41.4	31.1	35.5

‡ Reporting standards not met. Values suppressed to protect participant confidentiality.

¹In Florida, Treatment 1 is *Literacy Express*.

²In Florida, Treatment 2 is *DLM Early Childhood Express* supplemented with *Open Court Reading Pre-K*.

SOURCE: PCER Parent Interview (Fall 2003, Spring 2004, and Spring 2005).

Table 9.2. Primary caregiver demographic characteristics for *Literacy Express* and *DLM Early Childhood Express* supplemented with *Open Court Reading Pre-K*

Characteristics	Full sample n = 268	Curriculum comparison		
		Control n = 86	Treatment 1 ¹ n = 90	Treatment 2 ² n = 92
Age at baseline (years), mean	31.2	30.0	31.6	31.8
Marital status (%)				
Married	36.2	24.4	41.1	42.4
Separated/Divorced	20.1	15.1	21.1	23.9
Widowed	‡	‡	‡	0.0
Never Married	42.5	59.3	35.6	33.7
Race/ethnicity (%)				
White, non-Hispanic	34.5	25.9	41.6	35.5
African American, non-Hispanic	58.4	71.8	50.6	53.8
Hispanic	4.9	‡	4.5	8.6
Asian or Pacific Islander	‡	0.0	‡	0.0
Native American	‡	‡	0.0	0.0
Multiple/other	‡	0.0	‡	‡
Educational level (%)				
Did not finish high school	13.2	23.5	9.1	7.5
High school diploma or GED	34.2	38.8	36.4	28.0
Some college	37.6	28.2	42.0	42.0
College graduate	15.0	9.4	12.5	22.6
Employment (%)				
Full-time	63.2	59.3	67.8	62.4
Part-time	11.5	10.5	13.3	10.8
Unemployed	23.4	27.9	17.8	24.7
Other	1.9	‡	‡	‡

‡ Reporting standards not met. Values suppressed to protect participant confidentiality.

¹In Florida, Treatment 1 is *Literacy Express*.

²In Florida, Treatment 2 is *DLM Early Childhood Express* supplemented with *Open Court Reading Pre-K*.

SOURCE: PCER Parent Interview (Fall 2003, Spring 2004, and Spring 2005).

Table 9.3. Preschool teacher characteristics for *Literacy Express* and *DLM Early Childhood Express* supplemented with *Open Court Reading Pre-K*

Characteristics	Full sample n = 30	Curriculum comparison		
		Control n = 9	Treatment 1 ¹ n = 10	Treatment 2 ² n = 11
Gender (% female)	97.0	89.0	100.0	100.0
Race/ethnicity (%)				
White, non-Hispanic	83.0	89.0	80.0	82.0
African American, non-Hispanic	13.0	0.0	‡	‡
Hispanic	‡	‡	0.0	0.0
Asian or Pacific Islander	0.0	0.0	0.0	0.0
Native American	0.0	0.0	0.0	0.0
Multiple/other	0.0	0.0	0.0	0.0
Educational level (%)				
High school diploma or GED	13.0	0.0	0.0	36.0
Associate's degree	‡	0.0	‡	‡
Bachelor's degree	53.0	78.0	50.0	36.0
Graduate degree	27.0	‡	40.0	‡
Current teaching license/certificate (%)	80.0	100.0	90.0	55.0
Child Development Associate (CDA) (%)	23.0	‡	‡	45.0
State-awarded preschool certificate (%)	40.0	‡	50.0	45.0
No credential (%)	0.0	0.0	0.0	0.0
Years of teaching experience, overall (mean)	15.9	17.6	15.4	15.1
Years of preschool teaching experience (mean)	9.3	10.7	10.4	7.1

‡ Reporting standards not met. Values suppressed to protect participant confidentiality.

¹In Florida, Treatment 1 is *Literacy Express*.

²In Florida, Treatment 2 is *DLM Early Childhood Express* supplemented with *Open Court Reading Pre-K*.

SOURCE: PCER Preschool Teacher Survey (Fall 2003 and Spring 2004).

Random Assignment

The Florida-FSU research team recruited 17 schools to participate in the study. Schools were rank ordered according to their letter grade (A, B, C, or D) using Florida's school grading report⁵ for each school. It was important to consider school grade as a blocking variable, because the letter grades represent schools with percentages of students who are functioning at categorically different levels of academic achievement based on the Florida Department of Education's grading system. The sample of 17 schools included 11 grade A schools, one grade B school, three grade C schools, one grade D school, and one school for which a grade could not be determined. Within each letter grade ranking, the research team ranked each school by the average number of years of teaching experience that the teachers had. Once the list of 16 graded schools was rank-ordered, the schools were grouped into triplets, and within each triplet, the schools were randomly assigned (using the random function in Excel) to one of three conditions (*Literacy Express*, *DLM Early Childhood Express* supplemented with *Open Court Reading Pre-K*, or control). The ungraded school and one

⁵ All schools in Florida receive a grade based on the following: (1) percentage of students meeting high standards of the Florida Comprehensive Assessment Test (FCAT)-achievement scores of Level 3 or above; (2) percentage of students making learning gains; and (3) adequate progress of the lowest 25 percent of the students in the school. Each school receives a certain number of points for each of three categories. The points are summed to create a total score. The total score is converted into a letter: grade A (410 points or more), grade B (380 to 409 points), grade C (320 to 379), grade D (280 to 319), and grade F (less than 250). For a grade of A, 95 percent were tested and at least 50 percent of the lowest readers must have made gains in the current school year. For a grade of B or C at least 50 percent of the lowest readers must have made gains in one or two consecutive years. Information reported here is based on the school grade categories in use during the 2003-04 school year. Source: <http://schoolgrades.fldoc.org>.

additional school that was a late entry to the project were randomly assigned separately to the *DLM Early Childhood Express* supplemented with *Open Court Reading Pre-K* curriculum condition and the control condition, respectively. Schools were randomly assigned to condition. The number of pre-kindergarten classrooms in each school ranged from one to three. Slightly more than half of the schools (9 of 17) had two pre-kindergarten classrooms per schools. At schools where there were two or more treatment group classrooms assigned to either *Literacy Express* or *DLM Early Childhood Express* supplemented with *Open Court Reading Pre-K* curriculum conditions, one of the two treatment group classrooms at those schools was then randomly assigned to a mentoring condition.⁶

Contamination

Because school was the unit of random assignment, all participating teachers within each school used the same curriculum, thus reducing the likelihood of contamination across conditions.

Control Condition

For all classrooms in the control condition, the school district was responsible for providing teachers with *High/Scope* curriculum training. The training provided to teachers in the control condition included a week-long summer institute conducted by *High/Scope* trainers prior to the start of the project, additional training sessions throughout the school year conducted by both *High/Scope* personnel and district personnel, and classroom visits by the *High/Scope* trainer. The evaluation, however, was not intended to be an evaluation of the *High/Scope* curriculum.

Data Collection

Mathematica Policy Research, Inc. (MPR) collected the child, parent, teacher, and school data for the Florida-FSU site for all three waves of data collection. The fall assessment data collection window for child assessments ranged from September 30, 2003 to November 17, 2003. The average delay from the beginning of the treatment (i.e., start of the school year) to the beginning of the fall assessment window was 42 days. The spring pre-kindergarten window was April 19, 2004 to June 15, 2004, and the kindergarten follow-up window was April 4, 2005 to June 6, 2005.

Attrition

Seventeen schools were randomly assigned to one of two treatment conditions or to the control condition. All 17 schools remained in the study throughout the pre-kindergarten year.

For the child assessment, the fall 2003 response rate was 95 percent; the spring 2004 pre-kindergarten response rate was 96 percent; and the kindergarten follow-up response rate was 80 percent.

Implementation

The Florida-FSU research team provided training and support to the treatment group teachers who implemented *Literacy Express* or *DLM Early Childhood Express* supplemented with *Open Court Reading Pre-K*. Teachers and teachers' aides in the two treatment curriculum conditions were provided with all required materials and received direct training in the use of these curriculum materials. Curriculum training was provided to *Literacy Express* classroom teachers from July 28, 2003 to July 31, 2003, and for those implementing *DLM Early Childhood Express* supplemented with *Open Court Reading Pre-K*, from July 30, 2003 to August 4, 2003. In each of the training sessions, the first 2 days were spent in a workshop setting and the other 2 were used for team planning. The workshop training session familiarized teachers and their aides with

⁶ Assignment to the mentoring conditions is a feature of the researcher's complementary research study. This report does not present findings based on the mentoring group assignment.

the new curriculum materials and provided hands-on experience in leading activities from the curricula. The workshop sessions were videotaped for those who were unable to attend to view at later times, so that all faculty and staff involved in the project would receive the same training. Make-up training using the videotapes was offered throughout the year as new staff joined the schools. Throughout the school year, teachers and teachers' aides received additional professional development opportunities in the use of the treatment curricula and related topics. All treatment group teachers attended a 2-hour professional development meeting specific to their assigned curriculum every other month. Attendance was documented at these meetings.

At sites (schools) where two or more treatment group classrooms were assigned to either *Literacy Express* or *DLM Early Childhood Express supplemented with Open Court Reading Pre-K* curriculum conditions, one of the two treatment classrooms at each site was randomly assigned to a mentoring condition. Throughout the school year, teachers in the mentoring condition received visits in their classrooms from the project's mentor teacher. These visits lasted on average hours per week, for a monthly average of 8 to 10 hours per class. During the visits, the mentor teacher acted as a coach providing the teacher with the opportunity to engage in collegial conversation and receive extra support in the implementation of the curriculum. The mentor used a combination of techniques to provide support the teachers and teachers' aides. The techniques included demonstrations, feedback, and troubleshooting in the use of the appropriate curriculum. The mentor teacher also worked with individual students and groups of students who were not responding to the curricula, who were nonverbal, or who simply needed more intense intervention strategies to be successful in the curricula.

Site-specific curriculum fidelity observations were conducted in both treatment and control classrooms in February 2004, and April/May 2004. Observations in each classroom consisted of two observational rating systems (Early Language and Literacy Classroom Observation [ELLCO] and Center for Improving the Readiness of Children for Learning and Education [CIRCLE] teacher observation tool) as well as two specific fidelity measures for *Literacy Express* and *DLM Early Childhood Express supplemented with Open Court Reading Pre-K*. The observational rating systems were completed following a 2.5 to 3-hour period of observation. The fidelity measures were completed every 20 to 30 minutes during the observation period. For 25 percent of the classrooms, a second observer completed these same fidelity and global ratings to provide an estimate of the reliability of the classroom measurement.

Implementation Fidelity Ratings

Literacy Express

Each research team used a global fidelity measure to rate the overall fidelity with which the curricula were implemented in the preschool year of the project. A four-point scale ranging from "Not at All" (0) to "High" (3) was used to rate each treatment classroom. Researchers were asked to use their site-specific implementation and fidelity data to rate each treatment classroom on the global fidelity measure as High, Medium, Low, or Not at All. Researchers were also asked to provide a global rating for the control group curriculum. *Literacy Express* was rated in the high Medium range (2.5) on the global implementation fidelity measure. The control group curriculum was rated at the Medium level (2.0).

DLM Early Childhood Express supplemented with Open Court Reading Pre-K

Each research team used a global fidelity measure to rate the overall fidelity with which the curricula were implemented in the preschool year of the project. A four-point scale ranging from "Not at All" (0) to "High" (3) was used to rate each treatment classroom. Researchers were asked to use their site-specific implementation and fidelity data to rate each treatment classroom on the global fidelity measure as High, Medium, Low, or Not at All. Researchers were also asked to provide a global rating for the control group curriculum. *DLM Early Childhood Express supplemented with Open Court Reading Pre-K* was rated in the high Medium range (2.3) on the global implementation fidelity measure. The control group curriculum was rated at the Medium level (2.0).

Impact Analysis Results

Because the Florida State University researchers evaluated two curricula, we present the results first for *Literacy Express* and then for *DLM Early Childhood Express* supplemented with *Open Court Reading Pre-K*. For each curriculum, we begin with the analyses of the child-level measures (i.e., mathematics, reading, phonological awareness, and language assessments) and then present the analyses of the classroom observation data.

Literacy Express—Child Outcomes

The unadjusted mean scores for child-level measures are reported in table C-10a in appendix C. Covariate adjusted mean differences and standard errors are reported in table D-10a in appendix D. For all analyses of child-level measures, the following covariates were included (a) child's age, (b) gender, (c) race/ethnicity, (d) disability status as reported by parent, and (e) mother's education. The student-level effect sizes (ESs) are presented in table 9.4.

Mathematics assessments

We used repeated measures linear spline models to analyze the data from all three mathematics measures (Woodcock Johnson [WJ] Applied Problems, Child Math Assessment-Abbreviated [CMA-A] Composite Score, and Shape Composition). There were no statistically detectable differences on these measures for the fall assessment.

There were no statistically detectable differences between groups on any of these measures for the spring pre-kindergarten or spring kindergarten assessments.

Based on the analyses for the three mathematics measures, we conclude that *Literacy Express* did not have a statistically detectable effect on mathematics relative to the control condition.

Reading assessments

Data from the three reading measures (Test of Early Reading Ability [TERA], WJ Letter Word Identification, and WJ Spelling) were analyzed using repeated measures linear spline models. There were no statistically significant differences on the TERA and WJ Spelling test for the fall assessment. There was, however, a statistically significant difference (ESs = .44. $p < .05$) favoring the *Literacy Express* group on the WJ Letter Word Identification scale for the fall assessment. This difference could indicate the failure of randomization to achieve equivalent groups at the start of treatment or an early treatment effect. Additional analyses of these data are provided in appendix A.

There were no statistically detectable differences between groups on any of these measures for the spring pre-kindergarten or spring kindergarten assessments.

Based on the analyses for the three reading measures, we conclude that *Literacy Express* did not have a statistically detectable effect on reading relative to the control condition.

Phonological awareness

The phonological awareness measures were the Preschool Comprehensive Test of Phonological and Print Processing (Pre-CTOPPP), Elision subtest, and the Comprehensive Test of Phonological Processing (CTOPP), Kindergarten, Elision subtest. We conducted a repeated measures analysis on the Pre-CTOPPP fall and spring pre-kindergarten data. There was no statistically detectable difference on the Pre-CTOPPP for the fall assessment.

There was no statistically detectable difference on the Pre-CTOPPP for the spring pre-kindergarten assessment.

We analyzed the kindergarten CTOPP data using analysis of covariance (ANCOVA). For the ANCOVA analysis, the covariates were the (a) Pre-CTOPPP fall assessment score, (b) child's gender, (c) age, (d) race/ethnicity, (e) disability status as reported by parent, and (f) mother's education. There was no statistically detectable difference between groups for the CTOPP spring kindergarten assessment.

Based on the analyses of the Pre-CTOPPP and CTOPP, we conclude that *Literacy Express* did not have a statistically detectable effect on phonological awareness relative to the control condition.

Language assessments

Data from the two language measures (Peabody Picture Vocabulary Test [PPVT] and Test of Language Development [TOLD] Grammatical Understanding subtest) were analyzed using repeated measures linear spline models. There were no statistically detectable differences on these measures for the fall assessment.

In the spring of the pre-kindergarten and kindergarten years, there were no statistically detectable differences between groups on either measure.

Based on the analyses of the two language measures, we conclude that *Literacy Express* did not have a statistically detectable effect on language development relative to the control condition.

Behavioral outcomes

We conducted a repeated measures analysis for all three pre-kindergarten social behavioral measures (Social Skills Rating System [SSRS] Social Skills scale, SSRS Problem Behaviors scale, and Preschool Learning Behaviors Scale [PLBS]). The covariates were (a) child's age, (b) gender, (c) race/ethnicity, (d) disability status as reported by the parent, and (e) mother's education. There were no statistically detectable differences on the behavioral measures for the fall assessment.

For the spring pre-kindergarten assessment, there were no statistically detectable differences between groups on any of the behavior measures.

We analyzed the data from the kindergarten versions of the three behavioral measures (SSRS Social Skills scale, SSRS Problem Behaviors scale, and Learning Behaviors Scale [LBS]) using analysis of covariance (ANCOVA). For the ANCOVA analyses, the covariates included (a) the fall pre-kindergarten score of the pre-kindergarten version of the relevant test, along with (b) child's age, (c) gender, (d) disability status as reported by the parent, (e) race/ethnicity, and (f) mother's education.

There was no statistically detectable difference between groups on the SSRS Social Skills and Problem Behaviors scales.

We obtained a statistically reliable impact on the Learning Behaviors Scale ($ES_s = -.38, p < .05$), such that children in the *Literacy Express* classrooms exhibited weaker learning behaviors relative to students in the control condition for the spring kindergarten assessment, but not the spring pre-kindergarten assessment.

Based on the analyses of the three behavioral measures, we conclude that *Literacy Express* did not have a statistically detectable effect on children's social and learning behaviors relative to the control condition during pre-kindergarten or kindergarten assessments.

Literacy Express—Classroom Outcomes

The unadjusted mean scores for classroom measures are reported in table C-10b in appendix C. Covariate adjusted mean differences and standard errors are reported in table D-10b in appendix D. For all analyses of classroom measures, the following variables were included in the model as covariates: (a) teacher has a BA degree, (b) previous teaching experience, (c) child/adult ratio in classroom, (d) average class size, (e) city size, and (f) geographic site. The classroom-level effect sizes (ES_c) are presented in table 9.4.

Overall classroom environment

We conducted a repeated measures analysis on the Early Childhood Environment Rating Scale-Revised (ECERS-R). There was no statistically detectable difference between groups for the fall observation.

There was a statistically detectable difference between the *Literacy Express* classrooms and the control classrooms on the ECERS-R in spring of pre-kindergarten ($ES_c = 1.29, p < .05$). Treatment group classrooms received higher global quality ratings relative to the control group classrooms.

Based on the analysis of the ECERS-R, we conclude that *Literacy Express* had a positive effect on overall classroom quality relative to the control condition.

Teacher-child relationships

We obtained observations on the Arnett Detachment, Harshness, Permissiveness, and Positive Interactions scales in fall and spring of the pre-kindergarten year, and conducted repeated measures analyses. There were no statistically detectable differences on these measures for the fall observation.

There were no statistically detectable differences between groups on the Arnett Detachment, Harshness, Permissiveness, or Positive Interaction scales for the spring pre-kindergarten observation.

Based on the analyses of the four Arnett scales, we conclude that *Literacy Express* did not have a statistically detectable effect on teacher-child relationships relative to the control condition.

Classroom instruction

We obtained observations on classroom instruction in (a) early literacy (Teacher Behavior Rating Scale [TBRS] Print and Letter Knowledge and Written Expression scales), (b) phonological awareness (TBRS Phonological Awareness scale), (c) language (TBRS Book Reading and Oral Language scales), and (d) early mathematics (TBRS Math Concepts scale) for the spring pre-kindergarten assessment only. To analyze these data, ANCOVAs were conducted; the covariates were: (a) teacher has a BA degree, (b) previous teaching experience, (c) child/adult ratio in classroom, (d) average class size, (e) city size, and (f) geographic site.

There were no statistically detectable differences on the Book Reading, Print and Letter Knowledge, Written Expression, Oral Language, or Math Concepts scales. There was a statistically significant difference on the Phonological Awareness ($ES_c = 1.26, p < .05$) scale, indicating that the *Literacy Express* teachers provided more instruction in phonological awareness relative to the instruction provided in the control classrooms.

Based on the analyses of the TBRS Print and Letter Knowledge and Written Expression scales, we conclude that *Literacy Express* did not have a statistically detectable effect on early literacy instruction relative to the control condition.

Based on the analysis of the TBRS Phonological Awareness scale, we conclude that *Literacy Express* had a positive effect on instruction in phonological awareness relative to the control condition.

Based on the analysis of the TBRS Book Reading and Oral Language scales, we conclude that *Literacy Express* did not have a statistically detectable effect on language instruction relative to the control condition.

Based on the analysis of the TBRS Math Concepts scale, we conclude that *Literacy Express* did not have a statistically detectable effect on early mathematics instruction relative to the control condition.

Summary of Findings for *Literacy Express*

The findings for *Literacy Express* are summarized in table 9.4.

Table 9.4. Effect sizes for Literacy Express

Measure	Student-level effect sizes (ES _s)		
	RM analysis Spring Pre-K	RM analysis Spring K	ANCOVA Spring K
Mathematics			
WJ Applied Problems	.05	-.02	—
CMA-A Mathematics Composite	-.02	-.21	—
Shape Composition ¹	-.01	-.14	—
Reading			
TERA	.17	-.11	—
WJ Letter Word Identification	.30	.08	—
WJ Spelling	.05	.06	—
Phonological awareness			
Pre-CTOPPP/CTOPP	.14	†	.08
Language			
PPVT	.17	.16	—
TOLD	-.04	.10	—
Behavior			
SSRS Social Skills	-.06	†	-.37
SSRS Problem Behavior ²	-.31	†	.22
PLBS/LBS	.17	†	-.38*
Measure	Classroom-level effect sizes (ES _c)		
	RM analysis Spring Pre-K	ANCOVA Spring Pre-K	
Global classroom quality			
ECERS-R	1.29*	—	
Teacher-child interaction			
Arnett Detachment ³	-1.09	—	
Arnett Harshness ³	-.84	—	
Arnett Permissiveness ³	.51	—	
Arnett Positive Interactions	.56	—	
Teacher instructional practices ⁴			
TBRS Book Reading	†	.49	
TBRS Oral Language	†	.25	
TBRS Phonological Awareness	†	1.26*	
TBRS Print and Letter Knowledge	†	1.07	
TBRS Written Expression	†	-.03	
TBRS Math Concepts	†	-.12	

— Not available.

† Not applicable. Four of the kindergarten student-level measures were not on the same scale as the pre-kindergarten measures. The classroom-level data were only collected during the pre-kindergarten year of the study.

* $p < .05$ ¹ Building Blocks, Shape Composition task² Higher scores on this scale represent more negative child behaviors.³ Lower scores on this scale represent a more positive classroom environment.⁴ ANCOVA models for the TBRS measures did not include baseline pretest scores because TBRS data were only collected in spring of the pre-kindergarten year.

NOTE: RM: Repeated Measures

ANCOVA: Analysis of covariance

Significance indications (p -values) in the table refer to the tests of contrasts between intervention and control groups that underlie the effect sizes reported here. Refer to the glossary for abbreviations of the measures.

SOURCE: The Preschool Curriculum Evaluation Research (PCER) Study.

DLM Early Childhood Express supplemented with Open Court Reading Pre-K—Child Outcomes

The unadjusted mean scores for child-level measures are reported in table C-11a in appendix C. Covariate adjusted mean differences and standard errors are reported in table D-11a in appendix D. For all analyses of child-level measures, the following covariates were included: (a) child's age, (b) gender, (c) race/ethnicity, (d) disability status as reported by parent, and (e) mother's education. The student-level effect sizes (ES_s) are presented in table 9.5.

Mathematics assessments

We conducted repeated measures linear spline models to analyze the data from all three mathematics measures (WJ Applied Problems, CMA-A Composite Score, and Shape Composition). There were no statistically detectable differences between groups on these measures for the fall assessment.

There were statistically reliable mean differences in scores on WJ Applied Problems for the spring pre-kindergarten assessment (ES_s = .36, $p < .01$)⁷ and the spring kindergarten assessment (ES_s = .48, $p < .001$) favoring children in the *DLM Early Childhood Express with Open Court Reading Pre-K* classrooms. There were no statistically detectable differences on the other two mathematics measures.

Based on the analyses for the three mathematics measures, we conclude that *DLM Early Childhood Express supplemented with Open Court Reading Pre-K* did not have a statistically detectable effect on mathematics relative to the control condition.

Reading assessments

We conducted repeated measures linear spline models to analyze the data from all three reading assessments. There were no statistically significant differences on the fall assessment for the TERA and the WJ Spelling. However, there was a statistically significant difference for the fall assessment on the WJ Letter Word Identification test (ES_s = .41, $p < .05$; follow-up analyses for this finding are included in appendix A).

There were statistically reliable mean differences on all three reading measures favoring students in the *DLM Early Childhood Express supplemented with Open Court Reading Pre-K* classrooms in spring of pre-kindergarten: TERA (ES_s = .68, $p < .001$), WJ Letter Word (ES_s = .51, $p < .01$), and WJ Spelling (ES_s = .46, $p < .01$).

For the spring kindergarten assessment, statistically reliable differences were obtained on two of the three reading measures (TERA, ES_s = .76, $p < .01$; WJ Letter Word Identification, ES_s = .50, $p < .01$), indicating that the difference in spring of pre-kindergarten was sustained through spring of the following year. There was no statistically detectable difference in scores on the WJ Spelling.

Based on the analyses for the three reading measures, we conclude that *DLM Early Childhood Express supplemented with Open Court Reading Pre-K* improved young children's early reading skills relative to the control condition.

Phonological awareness

The phonological awareness measures were the Pre-CTOPPP, Elision subtest, and the CTOPP, Kindergarten, Elision subtest. We conducted a repeated measures analysis on the Pre-CTOPPP fall and spring pre-kindergarten data. There was no statistically detectable difference on the Pre-CTOPPP for the fall assessment.

There was a statistically reliable difference favoring the *DLM Early Childhood Express supplemented with Open Court Reading Pre-K* condition in the spring of pre-kindergarten (Pre-CTOPPP, ES_s = .32, $p < .05$).

⁷ Significance indications (p -values) in the text refer to the tests of contrasts between intervention and control groups that underlie the effect sizes reported here.

We analyzed the kindergarten CTOPP data using analysis of covariance (ANCOVA). For the ANCOVA analysis, the covariates were the (a) Pre-CTOPPP fall assessment score, (b) child's gender, (c) age, (d) race/ethnicity, (e) disability status as reported by parent, and (f) mother's education. There was a statistically significant difference between groups on the CTOPP in the spring of kindergarten favoring the *DLM Early Childhood Express with Open Court Reading Pre-K* classrooms ($ES_s = .38, p < .05$).

Based on the analyses of the Pre-CTOPPP and CTOPP, we conclude that *DLM Early Childhood Express supplemented with Open Court Reading Pre-K* improved phonological awareness relative to the control condition.

Language assessments

Data from the two language measures (Peabody Picture Vocabulary Test [PPVT] and Test of Language Development [TOLD] Grammatical Understanding subtest) were analyzed using repeated measures linear spline models. There were statistically reliable differences on the TOLD for the fall assessment ($ES_s = .38, p < .05$; follow-up analyses for this finding are included in appendix A).

In the spring of the pre-kindergarten year, there were statistically reliable mean differences in scores on both language measures (PPVT: $ES_s = .40, p < .05$; TOLD Grammatical Understanding: $ES_s = .40, p < .01$). These differences were sustained through spring of the following year (PPVT: $ES_s = .48, p < .01$; TOLD Grammatical Understanding subtest: $ES_s = .46, p < .01$).

Based on the analyses of the two language measures, we conclude that *DLM Early Childhood Express supplemented with Open Court Reading Pre-K* improved children's language development relative to the control condition.

Behavioral outcomes

We conducted a repeated measures analysis for all three pre-kindergarten social behavioral measures (SSRS Social Skills scale, SSRS Problem Behaviors scale, and Preschool Learning Behaviors Scale [PLBS]). The covariates were (a) child's age, (b) gender, (c) disability status as reported by the parent, (d) race/ethnicity, and (e) mother's education. There were no statistically detectable differences on these measures for the fall assessment. For the spring pre-kindergarten assessment, there were no statistically detectable differences on any of these measures.

We analyzed the data from the kindergarten versions of the three behavioral measures (SSRS Social Skills scale, SSRS Problem Behaviors scale, and Learning Behaviors Scale [LBS]) using analysis of covariance (ANCOVA). For the ANCOVA analyses, the covariates included (a) the fall pre-kindergarten score of the pre-kindergarten version of the relevant test, along with (b) child's age, (c) gender, (d) race/ethnicity, (e) disability status as reported by the parent, and (f) mother's education.

There were no statistically detectable differences between groups on any of these measures for the spring kindergarten assessment.

Based on the analyses of the three behavioral measures, we conclude that *DLM Early Childhood Express supplemented with Open Court Reading Pre-K* did not have a statistically detectable effect on children's social and learning behaviors relative to the control condition.

DLM Early Childhood Express supplemented with Open Court Reading Pre-K— Classroom Outcomes

The unadjusted mean scores for classroom measures are reported in table C-11b in appendix C. Covariate adjusted mean differences and standard errors are reported in table D-11b in appendix D. For all analyses of classroom measures, the following variables were included in the model as covariates: (a) teacher has a BA degree, (b) previous teaching experience, (c) child/adult ratio in classroom, (d) average class size, (e) city size, and (f) geographic site. The classroom-level effect sizes (ES_c) are presented in table 9.5.

Overall classroom environment

We conducted a repeated measures analysis on the ECERS-R. There was no statistically detectable difference between groups on the fall observation.

No statistically significant difference between groups was obtained for the spring pre-kindergarten assessment.

Based on the analysis of the ECERS-R, we conclude that *DLM Early Childhood Express supplemented with Open Court Reading Pre-K* did not have a statistically detectable effect on overall classroom quality relative to the control condition.

Teacher-child relationships

We obtained observations on the Arnett Detachment, Harshness, Permissiveness, and Positive Interactions scales in fall and spring of the pre-kindergarten year, and conducted repeated measures analyses. There were no statistically detectable differences on these measures for the fall observation.

There were no statistically significant differences between groups on the Arnett Detachment, Harshness, Permissiveness, or Positive Interaction scales on the spring of pre-kindergarten observation.

Based on the analyses of the four Arnett scales, we conclude that *DLM Early Childhood Express supplemented with Open Court Reading Pre-K* did not have a statistically detectable effect on teacher-child relationships relative to the control condition.

Classroom instruction

We obtained observations on classroom instruction in (a) early literacy (TBRS Print and Letter Knowledge and Written Expression scales); (b) phonological awareness (TBRS Phonological Awareness scale); (c) language (TBRS Book Reading and Oral Language scales); and (d) early mathematics (TBRS Math Concepts scale) in the spring of the pre-kindergarten year only. To analyze these data, ANCOVAs were conducted; the covariates were: (a) teacher has a BA degree, (b) previous teaching experience, (c) child/adult ratio in classroom, (d) average class size, (e) city size, and (f) geographic site.

There were no statistically detectable differences on the Book Reading, Print and Letter Knowledge, Written Expression, Oral Language, or Math Concepts scales. There was a statistically significant difference on the Phonological Awareness scale ($ES_c = 1.41, p < .05$), indicating that the *DLM Early Childhood Express supplemented with Open Court Reading Pre-K* teachers provided more instruction in phonological awareness relative to teachers in the control classrooms.

Based on the analyses of the TBRS Print and Letter Knowledge and Written Expression scales, we conclude that *DLM Early Childhood Express supplemented with Open Court Reading Pre-K* did not have a statistically detectable effect on early literacy instruction relative to the control condition.

Based on the analysis of the TBRS Phonological Awareness scale, we conclude that *DLM Early Childhood Express supplemented with Open Court Reading Pre-K* had a positive effect on instruction in phonological awareness relative to the control condition.

Based on the analysis of the TBRS Book Reading and Oral Language scales, we conclude that *DLM Early Childhood Express supplemented with Open Court Reading Pre-K* did not have a statistically detectable effect on language instruction relative to the control condition.

Based on the analysis of the TBRS Math Concepts scale, we conclude that *DLM Early Childhood Express supplemented with Open Court Reading Pre-K* did not have a statistically detectable effect on early mathematics instruction relative to the control condition.

Summary of Findings for *DLM Early Childhood Express supplemented with Open Court Reading Pre-K*

The findings for *DLM Early Childhood Express supplemented with Open Court Reading Pre-K* are summarized in table 9.5.

Table 9.5. Effect sizes for *DLM Early Childhood Express supplemented with Open Court Reading Pre-K*

Measure	Student-level effect sizes (ES _s)		
	RM analysis Spring Pre-K	RM analysis Spring K	ANCOVA Spring K
Mathematics			
WJ Applied Problems	.36**	.48***	—
CMA-A Mathematics Composite	.17	.13	—
Shape Composition ¹	.24	.09	—
Reading			
TERA	.68***	.76**	—
WJ Letter Word Identification	.51**	.50**	—
WJ Spelling	.46**	.22	—
Phonological awareness			
Pre-CTOPPP/CTOPP	.32*	†	.38*
Language			
PPVT	.40*	.48**	—
TOLD	.40**	.46**	—
Behavior			
SSRS Social Skills	-.11	†	-.18
SSRS Problem Behavior ²	.11	†	.01
PLBS/LBS	-.16	†	-.13
	Classroom-level effect sizes (ES _c)		
Measure	RM analysis Spring Pre-K	ANCOVA Spring Pre-K	
Global classroom quality			
ECERS-R	.34	—	
Teacher-child interaction			
Arnett Detachment ³	-.06	—	
Arnett Harshness ³	-.70	—	
Arnett Permissiveness ³	.05	—	
Arnett Positive Interactions	.43	—	
Teacher instructional practices ⁴			
TBRS Book Reading	†	.01	
TBRS Oral Language	†	-.33	
TBRS Phonological Awareness	†	1.41*	
TBRS Print and Letter Knowledge	†	.91	
TBRS Written Expression	†	-.58	
TBRS Math Concepts	†	-.46	

— Not available.

† Not applicable. Four of the kindergarten student-level measures were not on the same scale as the pre-kindergarten measures. The classroom-level data were only collected during the pre-kindergarten year of the study.

* $p < .05$; ** $p < .01$; *** $p < .001$

¹ Building Blocks, Shape Composition task

² Higher scores on this scale represent more negative child behaviors.

³ Lower scores on this scale represent a more positive classroom environment.

⁴ ANCOVA models for the TBRS measures did not include baseline pretest scores because TBRS data were only collected in spring of the pre-kindergarten year.

NOTE: RM: Repeated Measures

ANCOVA: Analysis of covariance

Significance indications (p -values) in the table refer to the tests of contrasts between intervention and control groups that underlie the effect sizes reported here. Refer to the glossary for abbreviations of the measures.

SOURCE: The Preschool Curriculum Evaluation Research (PCER) Study.

Chapter 10. *Pre-K Mathematics supplemented with DLM Early Childhood Express Math software:* **University of California, Berkeley/University at Buffalo, State University of New York (California/New York sites)**

Curriculum

The University of California, Berkeley, and the University at Buffalo, State University of New York (California/New York) research team implemented the *Pre-K Mathematics* curriculum supplemented with the *DLM Early Childhood Express Math software* in preschool classrooms in California and New York.

The *Pre-K Mathematics* curriculum consisted of 29 small-group mathematics activities with concrete manipulatives for use by teachers and children in preschool classrooms as well as 19 home mathematics activities and materials for use by parents and preschool-age children in home settings. The teacher's manual provided a curriculum plan that linked small-group classroom activities to home activities.

Teachers conducted small-group mathematics activities twice per week with all pre-kindergarten children. Small-group activities are conducted with groups of 4 to 6 children for approximately 20 minutes per group. Teachers completed Assessment Record Sheets specifically tied to the mathematics activity during each small-group session. In addition to these structured activities, similar mathematics materials and activities were available to children in classroom mathematics centers for use during free play. Materials for home mathematics activities were sent home every 1 to 2 weeks and corresponded conceptually to the classroom mathematics activities.

The *DLM Early Childhood Express Math software* included 26 numerical, quantitative, geometric, and spatial activities. The *DLM Early Childhood Express Math software* is a component of Building Blocks, a research-based mathematics curriculum that addresses (a) geometric and spatial ideas and skills and (b) numeric and quantitative ideas and skills. Working with the *DLM Early Childhood Express Math software*, children use pattern blocks and tangrams to complete puzzles.

The software program provided individualized pre-kindergarten mathematics instructional activities for children to use approximately twice a week. Curriculum implementation was conducted over a 36-week period. Activities were scheduled such that children engaged in conceptually related small-group, home, and computer mathematics activities during the same week. Teachers were encouraged to present information from Assessment Record Sheets and to discuss children's mathematics learning during routine parent-teacher conferences.

Sample

The California/New York research team recruited five Head Start and public pre-kindergarten programs in California and two Head Start and public pre-kindergarten programs in New York. A total of 40 teachers/classrooms (20 in each state) were recruited from these Head Start and public pre-kindergarten programs to participate in the study. Twenty-six (12 in California and 14 in New York) of the 40 classrooms were full-day pre-kindergarten programs. Consent letters were sent home to the parents of all eligible children in each classroom. Teachers and other classroom staff assisted with the recruitment of families. A sample of 316 children (159 treatment, 157 control) and parents were recruited for participation in the study. Data were collected on a total of 314 children and 263 parents at the time of the fall assessment.

The kindergarten sample included 309 children. Data were collected on 283 children and 246 parents at the time of the spring kindergarten assessment.

Children and Families

The children were 4.3 years of age at the time of the fall assessment data collection and almost half (48%) were male. The sample included African American (45%), Hispanic (23%), and White (18%) preschoolers. The racial/ethnic composition of the sample of children varied based on the geographic location of the sample. The California sample was primarily African American (48%) or Hispanic (35%). A larger percentage of White children (36%) were represented in the New York sample. Table 10.1 provides additional information on the demographic characteristics of the children in the California and New York study samples. At baseline, there were more boys in the control group classrooms relative to those assigned to the *Pre-K Mathematics supplemented with DLM Early Childhood Express Math software* classrooms (52% vs. 43%, $p = .05$).

The demographic characteristics of the primary caregivers, who were most often the biological or adoptive mother, are presented in table 10.2. The average age of the primary caregiver was 32 years. Almost half (43%) of the primary caregivers were never married; 40 percent were married at the time of the fall assessment data collection. Half of the primary caregivers reported having had some college education (40%) or a college degree (11%); 27 percent had a high school diploma or GED; and 23 percent had not finished high school. A large percentage (40%) of the primary caregivers were not working at the time of the fall data collection. Some were employed full-time (37%) or part-time (20%). There were no statistically detectable differences between the treatment and control groups on the primary caregiver characteristics.

Teachers

Forty teachers participated in the preschool year intervention study; all were female. The racial/ethnic composition of the sample included White (38%), African American (33%), Hispanic (13%), and Asian (10%) teachers. On average, the preschool teachers had 19 years of teaching experience, with an average of 12 years of experience teaching preschool. The majority of teachers had a bachelor's (33%) or graduate (40%) degree. An additional 18 percent had an associate's degree, and 10 percent had a high school diploma or GED. The majority of teachers reported having a current teaching license/certificate (78%); state-awarded preschool certificate (68%); or a Child Development Associate (CDA) credential (33%). Table 10.3 provides additional information on the characteristics of the preschool sample of teachers. There were no statistically detectable differences between the treatment and control groups on the teacher characteristics.

Programs/Classrooms

The average preschool class size was 22.4 children in California, and 14.4 children in New York. The child-staff ratio was on average 7 to 1 in California, and 6.7 to 1 in New York.

Table 10.1. Child demographic characteristics for *Pre-K Mathematics supplemented with DLM Early Childhood Express Math software*

Characteristics	Full sample n = 316	Curriculum comparison	
		Control n = 157	Treatment n = 159
<i>Pre-K Mathematics supplemented with DLM Early Childhood Express Math software: California and New York</i>			
Age at baseline (years), mean	4.3	44.4	4.4
Gender (% male)	47.5	52.2*	42.7
Race/ethnicity (%)			
White, non-Hispanic	17.9	13.8	21.9
African American, non-Hispanic	44.7	49.0	40.4
Hispanic	23.0	22.8	23.3
Asian or Pacific Islander	3.1	‡	5.5
Native American	‡	‡	‡
Multiple/other	10.3	13.1	7.5
Child disability status (parent reported, %)	10.3	9.6	11.0
Characteristics	Full sample n = 160	Curriculum comparison	
		Control n = 80	Treatment n = 80
<i>Pre-K Mathematics supplemented with DLM Early Childhood Express Math software: California</i>			
Age at baseline (years), mean	4.5	4.4	4.5
Gender (% male)	46.9	57.5	36.3
Race/ethnicity (%)			
White, non-Hispanic	‡	‡	‡
African American, non-Hispanic	47.7	41.0	54.7
Hispanic	34.6	38.5	30.7
Asian or Pacific Islander	5.9	‡	10.7
Native American	0.0	0.0	0.0
Multiple/other	10.5	17.9	‡
Child disability status (parent reported, %)	6.7	7.9	5.4
Characteristics	Full sample n = 154	Curriculum comparison	
		Control n = 77	Treatment n = 77
<i>Pre-K Mathematics supplemented with DLM Early Childhood Express Math software: New York</i>			
Age at baseline (years), mean	4.3	4.3	4.3
Gender (% male)	48.1	46.8	49.4
Race/ethnicity (%)			
White, non-Hispanic	36.2	28.4	43.7
African American, non-Hispanic	41.3	58.2	25.4
Hispanic	10.1	‡	15.5
Asian or Pacific Islander	0.0	0.0	0.0
Native American	‡	‡	‡
Multiple/other	10.1	7.5	12.7
Child disability status (parent reported, %)	15.3	12.2	17.7

‡ Reporting standards not met. Values suppressed to protect participant confidentiality.

* $p < .05$

SOURCE: PCER Parent Interview (Fall 2003, Spring 2004, and Spring 2005).

Table 10.2. Primary caregiver demographic characteristics for *Pre-K Mathematics supplemented with DLM Early Childhood Express Math software*

Characteristics	Full sample n = 261	Curriculum comparison	
		Control n = 125	Treatment n = 136
Age at baseline (years), mean	32.5	31.9	33.0
Marital status (%)			
Married	39.6	43.2	36.3
Separated/Divorced	15.0	12.8	17.0
Widowed	2.3	3.2	‡
Never Married	43.1	40.8	45.2
Race/ethnicity (%)			
White, non-Hispanic	21.7	18.5	24.6
African American, non-Hispanic	42.6	46.0	39.6
Hispanic	22.5	25.0	20.1
Asian or Pacific Islander	4.7	‡	8.2
Native American	‡	‡	‡
Multiple/other	7.4	8.9	6.0
Educational level (%)			
Did not finish high school	22.7	27.4	18.4
High school diploma or GED	26.9	30.6	23.5
Some college	39.6	33.1	45.6
College graduate	10.8	8.9	12.5
Employment (%)			
Full-time	37.2	32.8	41.2
Part-time	19.9	17.6	22.1
Unemployed	40.2	46.4	34.6
Other	2.7	3.2	‡

‡ Reporting standards not met. Values suppressed to protect participant confidentiality.

SOURCE: PCER Parent Interview (Fall 2003, Spring 2004, and Spring 2005).

Table 10.3. Preschool teacher characteristics for *Pre-K Mathematics supplemented with DLM Early Childhood Express Math software*

Characteristics	Full sample n = 40	Curriculum comparison	
		Control n = 20	Treatment n = 20
Gender (% female)	100.0	100.0	100.0
Race/ethnicity (%)			
White, non-Hispanic	38.0	40.0	37.0
African American, non-Hispanic	33.0	40.0	26.0
Hispanic	13.0	‡	‡
Asian or Pacific Islander	10.0	0.0	21.0
Native American	0.0	0.0	0.0
Multiple/other	‡	‡	0.0
Educational level (%)			
High school diploma or GED	10.0	‡	‡
Associate's degree	18.0	20.0	‡
Bachelor's degree	33.0	40.0	25.0
Graduate degree	40.0	35.0	45.0
Current teaching license/certificate (%)	78.0	80.0	75.0
Child Development Associate (CDA) (%)	33.0	35.0	32.0
State-awarded preschool certificate (%)	68.0	74.0	12.0
No credential (%)	‡	‡	0.0
Years of teaching experience, overall (mean)	19.0	19.5	18.5
Years of preschool teaching experience (mean)	12.4	13.4	11.4

‡ Reporting standards not met. Values suppressed to protect participant confidentiality.

SOURCE: PCER Preschool Teacher Survey (Fall 2003 and Spring 2004).

Random Assignment

A total of 40 Head Start and public preschools were randomly assigned in the fall of the pilot study year by the research team, using block randomization to either the treatment condition (*Pre-K Mathematics supplemented with DLM Early Childhood Express Math software*) or the control condition. Blocks were formed at the program level (five programs in California and two in New York), with teachers from Head Start and state-funded programs balanced by curriculum assignment in each site. In California, random assignment of classrooms to intervention and control conditions was done publicly in the presence of project staff and program staff for each of the five programs. The names of teachers who expressed a willingness to participate in the study were placed in a container and randomly drawn. The classroom of the first teacher whose name was drawn was assigned to the treatment condition. The classroom of the second teacher whose name was drawn was assigned to the control condition. This random assignment process continued until the designated number of classrooms had been assigned to each condition. To conduct a substudy of children from low-income Asian American families, two classrooms with large enrollments of Chinese American children were paired, such that when the teacher for one classroom was randomly assigned to a condition, the other classroom was automatically assigned to the other condition. Furthermore, two classrooms in which the language of instruction was Spanish were paired, such that when one classroom was randomly assigned to one condition (e.g., treatment) the other classroom was automatically assigned to the other condition (control). In New York, the school district and Head Start program administrators provided the research team with the names of teachers who were willing to participate in the study. Assignment to treatment and control conditions was then done publicly. The names were then randomly drawn, with the stipulation that there could not be a treatment and control teacher in the same building. A subsample of eight focal children was randomly selected in each classroom from the total number of consented children, balanced for age and gender. A total of 40 classrooms and 315 were recruited to participate in the study.

The same assignments were maintained for the second year (2003-04) of implementation as had been used during the pilot-study year (2002-03). In California, the research team was able to retain 8 of the 10 treatment classrooms and nine of the 10 control classrooms from the pilot year. They were able to retain 8 of the 10 treatment group teachers and 8 of the 10 control group teachers. Three preschool classrooms (two treatment and one control) were closed or converted by the program to serve age groups other than pre-kindergarten. Three replacement classrooms and their head teachers were added by randomly selecting them from the list of volunteers provided by the participating programs. In another control classroom, the teacher left after the pilot year and the program assigned a new teacher to the classroom. This classroom was retained in the control condition.

In New York, the research team was able to retain 8 of the 10 pilot-year treatment group teachers. They were able to retain 9 of the 10 control group teachers. Three replacement teachers (two in the treatment condition and one in the control condition) were added to the study sample. These three teachers were assigned to the existing treatment and control group classrooms (based on the initial intention that the classroom was the unit of random assignment). After the beginning of the school year, one of the treatment schools closed its pre-kindergarten program and the teacher, one of the replacements teachers, was reassigned. The district pre-kindergarten administrator was contacted and provided the research team with the original list of volunteers. A replacement teacher was randomly selected from this list.

Contamination

The California/New York researchers randomly assigned classrooms/teachers in each location to the intervention or control condition. To minimize the likelihood of contamination, intervention and control classrooms were located in different buildings. Furthermore, programs were asked to ensure that intervention and control teachers did not substitute in classrooms assigned to a condition different from their own. Finally, classroom observations using the Early Mathematics Classroom Observation instrument (described below), as well as periodic unannounced classroom visits in treatment and control classrooms by project staff did not reveal any evidence of contamination.

Control Condition

A number of curricula were represented in the control classrooms. Prevailing curricula included *Creative Curriculum*, *High Scope*, *Montessori*, specialized literacy curricula, and local school district and teacher-developed curricula. In New York, control group teachers in the public pre-kindergarten classrooms used the *BPS Benchmarks*, a curriculum that was developed by the local school district. Head Start classroom teachers in New York used a version of the *Creative Curriculum*.

Data Collection

RTI International (RTI) collected the child, teacher, and school data for the California and New York sites for all three waves of data collection. The California/New York research team was responsible for conducting the parent interviews in the preschool year, except for the few Chinese-speaking parents who were interviewed by a trained member of the grantee staff who spoke Chinese. In the kindergarten follow-up year, RTI staff completed the parent interviews. The fall assessment data collection window for child assessments ranged from September 22, 2003 to November 7, 2003. The average delay from the beginning of the treatment (i.e., start of the school year) to the beginning of the fall assessment window was 19 days in California and 14 days in New York. The spring pre-kindergarten window was April 7, 2004 to June 10, 2004, and the kindergarten follow-up window was May 2, 2005 to June 14, 2005.

The California/New York researchers supplemented the assessment of mathematics practices by intervention and control teachers by administering the Early Mathematics Classroom Observation (EMCO) (Klein and Starkey 2000). The EMCO measured the *amount* of classroom support for mathematical development by recording the number of children who participated in a mathematics activity and the duration of their

participation, thus yielding a measure of minutes-of-math support per child for a given classroom. This instrument also measured the *nature* of classroom support for mathematical development: (1) the conceptual domain supported, such as number and arithmetic or spatial and geometric knowledge; (2) the specific concepts and skill supported, such as cardinal number and counting sets of 1-10 objects; and (3) the general type of mathematics activity presented by the teacher (focal mathematics, which has a primarily mathematical goal, or embedded mathematics, such as a cooking activity, which includes some mathematics, but has a primarily nonmathematical goal).

The California/New York researchers also administered the Child Math Assessment (CMA) at pretest and posttest. The CMA is a comprehensive assessment of early mathematical knowledge. The assessment was comprised of 17 tasks, using concrete objects and encompassing a range of problem difficulty appropriate for pre-kindergarten children. This instrument assesses mathematical knowledge within several distinct areas, including number, arithmetic, space and geometry, measurement, and pattern knowledge.

Attrition

Forty classrooms were randomly assigned to treatment or control condition. All 40 classrooms remained in the study from the beginning of the pre-kindergarten year through the spring of the pre-kindergarten year.

For the child assessment, the fall assessment response rate was 99 percent, the spring 2004 pre-kindergarten response rate was 94 percent, and the kindergarten follow-up response rate was 92 percent.

Implementation

Most (16 of 20) teachers in the treatment condition were in their second year of implementation of the *Pre-K Mathematics supplemented with DLM Early Childhood Express Math software* at the time of the evaluation. The teacher sample included teachers who participated in the pilot year of the study (2002-03), and new teachers who started in 2003-04. During the pilot year, treatment group teachers in California participated in 4-day training workshops in late summer (September 10-13) and winter (February 3-6). In New York, treatment group teachers participated in 4-day training workshops in fall (September 26-27; October 3-4) and in winter (February 10-11 and 20-21). Ongoing on-site training was provided by project staff approximately twice per month, for an average of 17 on-site training sessions per teacher in California and 12.5 training sessions per teacher in New York. During the second year of the implementation, treatment group teachers attended a refresher workshop for 2 days in late summer. Throughout the preschool year, project staff observed and rated the implementation fidelity of small-group activities in each intervention classroom 1-2 times per month. Feedback was given to treatment group teachers at the end of those observation sessions. Staff members also observed teachers and children while they were using the computer-based mathematics activities, examined computer records of children's use of these activities, and provided feedback and training to teachers as needed.

The California/New York research team collected fidelity of implementation data, using the Fidelity of Implementation Record Sheet (Klein and Starkey 2002), as part of their formative evaluation of the mathematics curriculum. They collected data on the fidelity of implementation of small-group activities, computer activities, and home activities. Implementation fidelity data were collected in fall and spring of the pre-kindergarten year. The research team also administered the Early Mathematics Classroom Observation (Klein and Starkey 2000) to collect data on the amount of teacher-participant mathematics support per child that was provided in treatment and control classrooms, whether mathematics content was focal or embedded on other types of activities, and the conceptual breadth of mathematics support provided by teachers. For the home activities measure, parents were asked to report on how often mathematics activities were sent home, how often they used the activities, whether they liked the activities, whether the activities helped their children learn mathematics, and whether the activities gave them ideas about how to help their children learn mathematics.

On the site-specific fidelity measure, the overall fidelity scores for the small-group activities and computer mathematics activities were calculated by averaging scores from one fall and one spring fidelity observation. Small-group mathematics fidelity ranged from Moderate to High across classrooms, and computer mathematics fidelity ranged from low Moderate to High. Overall levels of fidelity were similar across the California (average = .87, Head Start classrooms; .92 state pre-kindergarten classrooms) and New York (average = .78, Head Start classrooms; .84 state pre-kindergarten classrooms) research sites.

Implementation Fidelity Ratings

Each research team used a global fidelity measure to rate the overall fidelity with which the curricula were implemented in the preschool year of the project. A four-point scale ranging from “Not at All” (0) to “High” (3) was used to rate each treatment classroom. Researchers were asked to use their site-specific implementation and fidelity data to rate each treatment classroom on the global fidelity measure as High, Medium, Low, or Not at All. Treatment implementation was rated between High and Medium in California (2.65) and New York (2.25) on the global fidelity measure. Researchers were also asked to provide a global rating for the control group classrooms. The fidelity of implementation ratings for the various control group curricula was at the Medium level (2.0) at both sites.

Impact Analysis Results

We begin with the analyses of the child-level measures (i.e., mathematics, reading, phonological awareness, and language assessments) and then present the analyses of the classroom observation data. Our discussion of the results focuses on the combined analysis for the California and New York sites.

Pre-K Mathematics supplemented with DLM Early Childhood Express Math software—Child Outcomes

The unadjusted mean scores for child-level measures are reported in table C-12a in appendix C. Covariate adjusted mean differences and standard errors are reported in table D-12a in appendix D. For all analyses of child-level measures, the following covariates were included: (a) child’s age, (b) gender, (c) race/ethnicity, (d) disability status as reported by parent, and (e) mother’s education. The student-level effect sizes (ES_s) are presented in table 10.4.

Mathematics assessments

We used repeated measures linear spline models to analyze the data from all three mathematics measures (Woodcock Johnson [WJ] Applied Problems, Child Math Assessment-Abbreviated [CMA-A] Composite Score, and Shape Composition). There were no statistically significant differences for the fall assessment for the WJ Applied Problems and the CMA-A Composite Score. However, there was a statistically significant difference for the fall assessment on the Shape Composition task (ES_s = .25, $p < .05$; follow-up analyses are included in appendix A).

There was no statistically detectable difference for the spring pre-kindergarten or kindergarten assessments on the WJ Applied Problems.

In spring of the pre-kindergarten year, there was a statistically reliable mean difference in scores on the CMA-A Composite Score (ES_s = .44, $p < .01$) favoring the *Pre-K Mathematics supplemented with DLM Early Childhood Express Math software* group. No difference was found for the CMA-A Composite Score for the spring kindergarten assessment.

In spring of the pre-kindergarten year, there was a statistically reliable mean difference in scores on Shape Composition (ES_s = .96, $p < .001$) favoring the *Pre-K Mathematics supplemented with DLM Early Childhood Express Math software* group. The advantage of the treatment group was maintained through spring of the kindergarten year (ES_s = .41, $p < .001$).

Based on the analyses for the three mathematics measures, we conclude that *Pre-K Mathematics supplemented with DLM Early Childhood Express Math software* had a positive effect on children's early mathematics skills at the end of pre-kindergarten relative to the control condition. *Pre-K Mathematics supplemented with DLM Early Childhood Express Math software* did not have a statistically detectable effect on children's mathematics skills at the end of kindergarten.

Reading assessments

Data from the three reading measures (Test of Early Reading Ability [TERA], WJ Letter Word Identification, and WJ Spelling) were analyzed using repeated measures linear spline models. There were no statistically detectable differences on any of the three reading measures at the fall assessment.

There were no statistically detectable differences on any of these measures for the spring pre-kindergarten or spring kindergarten assessments.

Based on the analyses of the three reading measures, we conclude that *Pre-K Mathematics supplemented with DLM Early Childhood Express Math software* did not have a statistically detectable effect on reading relative to the control condition.

Phonological awareness

The phonological awareness measures were the Preschool Comprehensive Test of Phonological and Print Processing (Pre-CTOPPP), Elision subtest, and the Comprehensive Test of Phonological Processing (CTOPP), Kindergarten, Elision subtest. We conducted a repeated measures analysis on the Pre-CTOPPP fall and spring pre-kindergarten data. There was no statistically detectable difference on the Pre-CTOPPP for the fall assessment.

There was no statistically detectable difference on the Pre-CTOPPP for the spring pre-kindergarten assessment.

We analyzed the kindergarten CTOPP data using analysis of covariance (ANCOVA). For the ANCOVA analysis, the covariates were the (a) Pre-CTOPPP fall assessment score, (b) child's gender, (c) age, (d) race/ethnicity, (e) disability status as reported by parent, and (f) mother's education. There was no statistically detectable difference between groups on the CTOPP for the spring kindergarten assessment.

Based on the analyses of the Pre-CTOPPP and CTOPP, we conclude that *Pre-K Mathematics supplemented with DLM Early Childhood Express Math software* did not have a statistically detectable effect on phonological awareness relative to the control condition.

Language assessments

Data from the two language measures (Peabody Picture Vocabulary Test [PPVT], Test of Language Development [TOLD] Grammatical Understanding subtest) were analyzed using repeated measures linear spline models. There were no statistically detectable differences at the fall assessment.

In the spring pre-kindergarten and kindergarten years, there were no statistically detectable differences between groups on either measure.

Based on the analyses of the two language measures, we conclude that *Pre-K Mathematics supplemented with DLM Early Childhood Express Math software* did not have a statistically detectable effect on language development relative to the control condition.

Behavioral outcomes

We conducted a repeated measures analysis for all three pre-kindergarten social behavioral measures (Social Skills Rating System [SSRS] Social Skills scale, SSRS Problem Behaviors scale, and Preschool Learning Behaviors Scale [PLBS]). The covariates were child's age, gender, disability status as reported by the parent, race/ethnicity, and mother's education. There were no statistically detectable differences on these measures for the fall assessment.

For the spring pre-kindergarten assessment, there were no statistically detectable differences on any of these measures.

We analyzed the data from the kindergarten versions of the three behavioral measures (SSRS Social Skills scale, SSRS Problem Behaviors scale, and Learning Behaviors Scale [LBS]) using analysis of covariance (ANCOVA). For the ANCOVA analyses, the covariates included (a) the fall pre-kindergarten score of the pre-kindergarten version of the relevant test, along with (b) child's age, (c) gender, (d) race/ethnicity, (e) disability status as reported by the parent, and (f) mother's education.

There were no statistically detectable differences between groups on any of these measures for the spring kindergarten assessment.

Based on the analyses of the three behavioral measures, we conclude that *Pre-K Mathematics supplemented with DLM Early Childhood Express Math software* did not have a statistically detectable effect on children's social and learning behaviors relative to the control condition.

Pre-K Mathematics supplemented with DLM Early Childhood Express Math software—Classroom Outcomes

The unadjusted mean scores for classroom measures are reported in table C-12b in appendix C. Covariate adjusted mean differences and standard errors are reported in table D-12b in appendix D. For all analyses of classroom measures, the following variables were included in the model as covariates: (a) teacher has a BA degree, (b) previous teaching experience, (c) child/adult ratio in classroom, (d) average class size, (e) city size, and (f) geographic site. The classroom-level effect sizes (ES_c) are presented in table 10.4.

Overall classroom environment

We conducted a repeated measures analysis on the Early Childhood Environment Rating Scale-Revised (ECERS-R). There was no statistically detectable difference between groups on the fall observation.

No statistically detectable difference between groups was obtained for the spring pre-kindergarten observation.

Based on the analysis of the ECERS-R, we conclude that *Pre-K Mathematics supplemented with DLM Early Childhood Express Math software* did not have a statistically detectable effect on overall classroom quality relative to the control condition.

Teacher-child relationships

We obtained observations on the Arnett Detachment, Harshness, Permissiveness, and Positive Interactions scales in fall and spring of the pre-kindergarten year, and conducted repeated measures analyses. There were no statistically detectable differences on these measures for the fall observation.

There were no statistically detectable differences between groups on the Arnett Detachment, Harshness, Permissiveness, or Positive Interaction scales in spring of the pre-kindergarten year.

Based on the analyses of the four Arnett scales, we conclude that *Pre-K Mathematics supplemented with DLM Early Childhood Express Math software* did not have a statistically detectable effect on teacher-child relationships relative to the control condition.

Classroom instruction

We obtained observations on classroom instruction in (a) early literacy (Teacher Behavior Rating Scale [TBRS] Print and Letter Knowledge and Written Expression scales); (b) phonological awareness (TBRS Phonological Awareness scale); (c) language (TBRS Book Reading and Oral Language scale); and (d) early mathematics (TBRS Math Concepts scale) in spring of pre-kindergarten only. To analyze these data, ANCOVAs were conducted. The covariates were: (a) teacher has a BA degree, (b) previous teaching experience, (c) child/adult ratio in classroom, (d) average class size, (e) city size, and (f) geographic site. The effect sizes are presented in table 10.4.

There were no statistically detectable differences on the TBRS Book Reading, Print and Letter Knowledge, Written Expression, Phonological Awareness, Oral Language, or Math Concepts scales.

Based on the analyses of the TBRS Print and Letter Knowledge and Written Expression scales, we conclude that *Pre-K Mathematics supplemented with DLM Early Childhood Express Math software* did not have a statistically detectable effect on early literacy instruction relative to the control condition.

Based on the analysis of the TBRS Phonological Awareness scale, we conclude that *Pre-K Mathematics supplemented with DLM Early Childhood Express Math software* did not have a statistically detectable effect on instruction in phonological awareness relative to the control condition.

Based on the analysis of the TBRS Book Reading and Oral Language scales, we conclude that *Pre-K Mathematics supplemented with DLM Early Childhood Express Math software* did not have a statistically detectable effect on language instruction relative to the control condition.

Based on the analysis of the TBRS Math Concepts scale, we conclude that *Pre-K Mathematics supplemented with DLM Early Childhood Express Math software* did not have a statistically detectable effect on early mathematics instruction relative to the control condition.

Summary of Findings for *Pre-K Mathematics supplemented with DLM Early Childhood Express Math software*

The findings for *Pre-K Mathematics supplemented with DLM Early Childhood Express Math software* are summarized in table 10.4.

Table 10.4. Effect sizes for *Pre-K Mathematics supplemented with DLM Early Childhood Express Math software*

Measure	Student-level effect sizes (ES _s)		
	RM analysis Spring Pre-K	RM analysis Spring K	ANCOVA Spring K
Mathematics			
WJ Applied Problems	.22	.13	—
CMA-A Mathematics Composite	.44**	.13	—
Shape Composition ¹	.96***	.41***	—
Reading			
TERA	.13	.31	—
WJ Identification	-.01	.22	—
WJ Spelling	.20	.03	—
Phonological awareness			
Pre-CTOPPP/CTOPP	.04	†	-.11
Language			
PPVT	.17	.11	—
TOLD	.17	.08	—
Behavior			
SSRS Social Skills	.22	†	.06
SSRS Problem Behavior ²	-.09	†	-.01
PLBS/LBS	.09	†	.01
Measure	Classroom-level effect sizes (ES _c)		
	RM analysis Spring Pre-K	ANCOVA Spring Pre-K	
Global classroom quality			
ECERS-R	.05	—	
Teacher-child interaction			
Arnett Detachment ³	-.37	—	
Arnett Harshness ³	.18	—	
Arnett Permissiveness ³	-.45	—	
Arnett Positive Interactions	.16	—	
Teacher instructional practices ⁴			
TBRS Book Reading	†	.07	
TBRS Oral Language	†	.19	
TBRS Phonological Awareness	†	.38	
TBRS Print and Letter Knowledge	†	.07	
TBRS Written Expression	†	-.12	
TBRS Math Concepts	†	.57	

— Not available.

† Not applicable. Four of the kindergarten student-level measures were not on the same scale as the pre-kindergarten measures. The classroom-level data were only collected during the pre-kindergarten year of the study.

** $p < .01$; *** $p < .001$

¹ Building Blocks, Shape Composition task

² Higher scores on this scale represent more negative child behaviors.

³ Lower scores on this scale represent a more positive classroom environment.

⁴ ANCOVA models for the TBRS measures did not include baseline pretest scores because TBRS data were only collected in spring of the pre-kindergarten year.

NOTE: RM: Repeated Measures

ANCOVA: Analysis of covariance

Significance indications (p -values) in the table refer to the tests of contrasts between intervention and control groups that underlie the effect sizes reported here. Refer to the glossary for abbreviations of the measures.

SOURCE: The Preschool Curriculum Evaluation Research (PCER) Study.

Chapter 11. *Project Approach*: Purdue University and University of Wisconsin-Milwaukee (Wisconsin site)

Curriculum

The Purdue University and University of Wisconsin-Milwaukee (Purdue/Wisconsin) research team implemented the *Project Approach* curriculum. *Project Approach* is a set of teaching strategies that enables teachers to guide children through in-depth investigations of real world topics. The curriculum is designed to use children's interests as the starting point for organizing and developing classroom learning activities. There are three curriculum components that address children's learning needs: spontaneous play, systematic instruction, and project work.

A *project* is defined as an in-depth study of a real world topic that is worthy of children's attention and effort. Projects can be incorporated into an existing classroom instructional program and can extend over several days or weeks. The structural features of the *Project Approach* include discussion, fieldwork, representation, investigation, and display. During the preliminary planning stage, the teacher selects the topic of study (based primarily on classroom learning goals, children's interests, and the availability of local resources). The teacher then brainstorms her own experience, knowledge, and ideas and represents them in a topic web. This topic web is revised throughout the project and used for recording the progress of the project. In *Project Approach* classrooms, the daily schedule is to be structured so that children and teachers spend at least 45 to 60 minutes engaged in investigation and discovery, typically in small groups.

Sample

The Purdue/Wisconsin research team recruited public pre-kindergarten classrooms for participation in the study. The research team recruited 13 teachers from 12 different schools. The recruitment of parents and children began at the start of the preschool year and continued through the first 6 weeks of school. Teachers assisted with the recruitment of families. Parents were offered an incentive for completing the parent interviews. A sample of 204 children (114 treatment, 90 control) and parents were recruited for participation in the study. Data were collected on 204 children and 176 parents at the time of the fall baseline data collection.

In the follow-up year of the evaluation, the sample of schools went from 12 in pre-kindergarten to more than 37 in kindergarten. The sample of classrooms went from 13 preschool to 58 kindergarten classrooms. The kindergarten sample included 156 children and 153 parents from the original sample of participants. Data were collected on 150 children and 122 parents.

Children and Families

The children were 4.6 years of age at the time of baseline data collection and slightly more than half (53%) were male. The racial/ethnic composition of the sample was diverse: African American (40%), White (28%), and Hispanic (17%). Table 11.1 provides additional information on the demographic characteristics of the children in the study sample. At baseline, the treatment group had a higher percentage of African American children relative to the control group (52% vs. 24%, $p < .01$).

The demographic characteristics of the primary caregivers, who were most often the biological or adoptive mother, are presented in table 11.2. The average age of the primary caregiver was 31.7 years. Half (50%) the primary caregivers were married and 38 percent were never married. Half (51%) reported having had some college or a college degree; 32 percent had a high school diploma or GED; and 17 percent had not finished

Table 11.1. Child demographic characteristics for *Project Approach*

Characteristics	Full sample n = 204	Curriculum comparison	
		Control n = 90	Treatment n = 114
Age at baseline (years), mean	4.6	4.6	4.6
Gender (% male)	52.9	56.7	50.0
Race/ethnicity (%)			
White, non-Hispanic	28.2	36.6	21.2
African American, non-Hispanic	39.8	24.4	52.5**
Hispanic	17.1	20.7	14.1
Asian or Pacific Islander	‡	0.0	‡
Native American	‡	0.0	‡
Multiple/other	13.3	18.3	9.1
Child disability status (parent reported, %)	17.7	16.5	18.8

‡ Reporting standards not met. Values suppressed to protect participant confidentiality.

** $p < .01$

SOURCE: PCER Parent Interview (Fall 2003, Spring 2004, and Spring 2005).

Table 11.2. Primary caregiver demographic characteristics for *Project Approach*

Characteristics	Full sample n = 176	Curriculum comparison	
		Control n = 80	Treatment n = 96
Age at baseline (years), mean	31.7	32.0	31.4
Marital status (%)			
Married	50.0	63.8*	38.5
Separated/Divorced	11.4	11.3	11.5
Widowed	‡	0.0	‡
Never Married	38.1	25.0	49.0
Race/ethnicity (%)			
White, non-Hispanic	40.0	51.9	30.2
African American, non-Hispanic	41.7	25.3	55.2
Hispanic	9.1	12.7	6.3
Asian or Pacific Islander	‡	0.0	‡
Native American	2.3	‡	‡
Multiple/other	6.3	7.6	5.2
Educational level (%)			
Did not finish high school	17.2	14.1	19.8
High school diploma or GED	32.2	24.4	38.5
Some college	29.9	35.9	25.0
College graduate	20.7	25.6	16.7
Employment (%)			
Full-time	52.3	51.3	53.1
Part-time	18.8	25.0	13.5
Unemployed	27.8	21.3	33.3
Other	‡	‡	0.0

‡ Reporting standards not met. Values suppressed to protect participant confidentiality.

* $p < .05$

SOURCE: PCER Parent Interview (Fall 2003, Spring 2004, and Spring 2005).

high school. More than half (52%) of the primary caregivers were employed full-time, 19 percent were employed part-time, and 27.8 percent were unemployed. At baseline, a higher percentage of parents were married in the control curriculum relative to those assigned to *Project Approach* (64% vs. 39%, $p < .05$).

Teachers

There were 12 teachers who participated in the preschool year intervention study. All of the preschool teachers were female, and all were White. On average, the preschool teachers had 11 years of teaching experience, with an average of 8 years of experience teaching preschool. All of the teachers had a bachelor's (54%) or graduate (46%) degree, and all reported having a current teaching license/certificate. Table 11.3 provides additional information on the characteristics of the preschool sample of teachers. At baseline, a higher percentage of teachers in the control had more years of teaching experience (17 vs. 6, $p < .01$), and years of preschool teaching experience (12 vs. 5, $p < .05$) relative to those assigned to *Project Approach*.

Table 11.3. Preschool teacher characteristics for *Project Approach*

Characteristics	Full sample n = 13	Curriculum comparison	
		Control n = 6	Treatment n = 7
Gender (% female)	100.0	100.0	100.0
Race/ethnicity (%)			
White, non-Hispanic	100.0	100.0	100.0
African American, non-Hispanic	0.0	0.0	0.0
Hispanic	0.0	0.0	0.0
Asian or Pacific Islander	0.0	0.0	0.0
Native American	0.0	0.0	0.0
Multiple/other	0.0	0.0	0.0
Educational level (%)			
High school diploma or GED	0.0	0.0	0.0
Associate's degree	0.0	0.0	0.0
Bachelor's degree	54.0	‡	71.0
Graduate degree	46.0	67.0	‡
Current teaching license/certificate (%)	100.0	100.0	100.0
Child Development Associate (CDA) (%)	‡	0.0	‡
State-awarded preschool certificate (%)	36.0	‡	‡
No credential (%)	0.0	0.0	0.0
Years of teaching experience, overall (mean)	11.5	17.3**	6.4
Years of preschool teaching experience (mean)	8.4	12.3*	5.0

‡ Reporting standards not met. Values suppressed to protect participant confidentiality.

* $p < .05$; ** $p < .01$

SOURCE: PCER Preschool Teacher Survey (Fall 2003 and Spring 2004).

Programs/Classrooms

The average preschool class size was 24.2 children. The child-staff ratio was on average 15.2 children to one teacher or program staff person.

Random Assignment

Randomization was done during the pilot year of curriculum implementation. Twelve of 57 eligible teachers agreed to participate in the study. They were randomly assigned to the treatment and control conditions after stratifying for racial/ethnic composition of families served by the schools. However, because of teacher attrition between the pilot year and the second year of implementation, the pilot-year randomization was not maintained in year 2. In the second year of implementation, the district administrator provided the research team with a list of eligible schools from which to recruit a study sample for the second year of the study. Two pilot-year control group teachers and 11 newly recruited teachers were randomly assigned to 7 treatment and 6 control classrooms. The Purdue/Wisconsin research team randomly assigned 13 teachers to the experimental conditions (7 treatment and 6 control classrooms). The names of the teachers were placed in a container and randomly drawn and assigned to either the treatment or control group. In all schools but one, there was only one preschool classroom. In one school with two classrooms, both classrooms/teachers were assigned to the same condition (the treatment group) to avoid contamination. In all other schools, only one teacher/classroom was assigned to either the treatment or control condition. A total of 13 classrooms and 204 children took part in the study.

Contamination

Because all of the classrooms in each school were assigned to either the treatment or the control group, there was little risk of contamination across the treatment and control conditions.

Control Condition

The school district provided all preschool classrooms with the *Doors to Discovery* and *Growing with Mathematics* curriculum materials but these curricula were not used consistently across all of the classrooms. In the control classrooms, teachers reported implementing their own teacher-developed, nonspecific curricula when the research team asked them to report on the curriculum in use.

Data Collection

RT International (RTI) collected the child, teacher, and school data for the Wisconsin site for all three waves of data collection. The Purdue/Wisconsin research team was responsible for conducting the parent interviews in the preschool year. In the kindergarten follow-up year, RTI staff completed the parent interviews. The fall assessment data collection window for child assessments ranged from September 15, 2003 to October 31, 2003. The average delay from the beginning of the treatment (i.e., start of the school year) to the beginning of the fall assessment window was 13 days. The spring pre-kindergarten window was April 11, 2004 to June 10, 2004, and the kindergarten follow-up window was April 2, 2005 to June 6, 2005.

Attrition

Thirteen classrooms were randomly assigned to treatment or control condition. All 13 classrooms remained in the study from the beginning of the pre-kindergarten year through the spring of the pre-kindergarten year.

For the child assessment, the baseline (fall, 2003) response rate was 100 percent; the spring 2004 response rate was 94 percent; and the kindergarten follow-up response rate was 96 percent.

Implementation

The research team provided training and support to the treatment group teachers to implement the *Project Approach* curriculum. On average, each treatment group teacher received 48 hours of training and individualized support during the 2003-04 preschool year (October 2003 through May 2004). The training and support activities included 18 hours of introductory training; 12 hours of follow-up training; and an

average of 12 hours of individual consultation time with the curriculum mentor during the mentor's regular visits to intervention classrooms.

The introductory training was held for 3 days (6 hours per day) at the beginning of the preschool year. The training workshop was conducted by an expert on the *Project Approach* and a co-author of *Young Investigators: The Project Approach in the Early Years*, the primary source of information on the *Project Approach* (Helm and Katz 2001). The content of the introductory training focused on benefits of the *Project Approach*; distinctions between projects and themes; criteria for selecting a good project topic; a detailed examination of the three phases of projects; and the use of webbing to link curriculum goals to project work. On the final day of the introductory training, participants visited a classroom in a local school (ineligible for the study) with a curriculum that included projects.

The follow-up training was held for 2 days in January 2004 (6 hours per day). The training began with a visit to study participant who demonstrated a high level of *Project Approach* implementation. The other treatment teachers toured her classroom, heard a presentation on the projects done to date in the classroom, and participated in a focused review of how various project activities were connected to curriculum goals for 4-year-old children. One goal of the visit was to help teachers strengthen integration of the *Project Approach* with other curriculum activities, including reading/writing, mathematics, and science. The follow-up training also included a presentation and critique of recent and/or on-going projects by each teacher, and a problem-solving discussion focused on challenges and barriers encountered by each teacher in implementing projects. In the final session of the follow-up training, each teacher generated an anticipatory planning web for the next project in her classroom, and received feedback and guidance from the trainer and other teachers.

In addition to group training sessions, the curriculum mentor conducted an average of 20.7 curriculum-related visits to each treatment classroom from October to May of the preschool year. On average, each mentoring visit was 2.8 hours in length, with 21 minutes of this time devoted to individualized consultation with the teacher about curriculum implementation. The rest of the mentor's time for each visit was devoted to classroom observation. The content of the mentoring visits focused primarily on: clarifications and reminders regarding components of the *Project Approach*; suggestions and feedback regarding planning and/or implementing project work (e.g., suggestions for experts and field visits); and provisions of resources to support project work (e.g., pizza recipes for children to use in a pizza project). The mentor completed a form for each visit that documented the observation of project-related displays and activities; specific feedback provided to the teacher based on the observation; suggestions offered to the teacher; resources provided by the mentor to the classroom; and specific actions the teacher agreed to take.

Site-specific curriculum fidelity data were collected three times in each of the seven classrooms implementing the *Project Approach*. The observations were conducted at three time points across the school year (November 2003 to January 2004; February to March 2004; and April to May 2004). Curriculum fidelity was measured with an observation and interview protocol. An experienced early childhood educator was trained by the project Principal Investigator to conduct the observations. The measure included items to address 12 main components: number of different types of displays; engaging, accessible displays; supports for project work; frequency of project work; level of engagement in project work; level of engagement in project-related work during free play/work time; level of child interest in project topic; use of experts and field visits; number and frequency of activities and materials; number and frequency of mathematics-related experiences; extent of opportunities for parent involvement; and teacher planning and documentation. The research team did not conduct curriculum fidelity observations in the control classrooms because of the wide variability in the use of curriculum materials in these classrooms. The school district provided all preschool classrooms with the *Doors to Discovery* and *Growing with Mathematics* curriculum materials, but these curricula were not used consistently across all of the classrooms. The research team collected observational and interview data on control classroom teachers' use of themes as part of their complementary research.

Implementation Fidelity Ratings

Each research team used a global fidelity measure to rate the overall fidelity with which the curricula were implemented in the preschool year of the project. A four-point scale ranging from “Not at All” (0) to “High” (3) was used to rate each treatment classroom. Researchers were asked to use their site-specific implementation and fidelity data to rate each treatment classroom on the global fidelity measure as High, Medium, Low, or Not at All. Researchers were also asked to provide a global rating for the control group curriculum. *Project Approach* was rated Medium (1.86) on the global implementation fidelity measure. The control group curriculum was also rated at the Medium level (2.00).

Impact Analysis

We begin with the analyses of the child-level measures (i.e., mathematics, reading, phonological awareness, language, and behavioral assessments) followed by classroom measures.

Project Approach—Child Outcomes

The unadjusted mean scores for child-level measures are reported in table C-13a in appendix C. Covariate adjusted mean differences and standard errors are reported in table D-13a in appendix D. For all analyses of child-level measures, the following covariates were included: (a) child’s age, (b) gender, (c) race/ethnicity, (d) disability status as reported by parent, and (e) mother’s education. The student-level effect sizes (ESs) are presented in table 11.4.

Mathematics assessments

We used repeated measures linear spline models to analyze the data from all three mathematics measures (Woodcock Johnson [WJ] Applied Problems, Child Math Assessment-Abbreviated [CMA-A] Composite Score, and Shape Composition). There were no statistically detectable differences on these measures for the fall assessment.

There were no statistically detectable differences between groups on the spring pre-kindergarten or kindergarten assessments on any of the mathematics assessments.

Based on the analyses for the three mathematics measures, we conclude that *Project Approach* did not have a statistically detectable effect on mathematics relative to the control condition.

Reading assessments

Data from the three reading measures (Test of Early Reading Ability [TERA], WJ Letter Word Identification, and WJ Spelling) were analyzed using repeated measures linear spline models. There were no statistically detectable differences on these measures for the fall assessment.

There were no statistically detectable differences on any of these measures for the spring pre-kindergarten or spring kindergarten assessments.

Based on the analyses for the three reading measures, we conclude that *Project Approach* did not have a statistically detectable effect on reading relative to the control condition.

Phonological awareness

The phonological awareness measures were the Preschool Comprehensive Test of Phonological and Print Processing (Pre-CTOPPP), Elision subtest, and the Comprehensive Test of Phonological Processing (CTOPP), Kindergarten, Elision subtest. We conducted a repeated measures analysis on the Pre-CTOPPP fall and spring pre-kindergarten data. There was no statistically detectable difference on the Pre-CTOPPP for the fall assessment.

There was no statistically detectable difference on the Pre-CTOPPP for the spring pre-kindergarten assessment.

We analyzed the kindergarten CTOPP data using analysis of covariance (ANCOVA). For the ANCOVA analysis, the covariates were the (a) Pre-CTOPPP fall assessment score, (b) child's gender, (c) age, (d) race/ethnicity, (e) disability status as reported by parent, and (f) mother's education. There was no statistically detectable difference between groups on the CTOPP for the spring kindergarten assessment.

Based on the analyses of the Pre-CTOPPP and CTOPP, we conclude that *Project Approach* did not have a statistically detectable effect on phonological awareness relative to the control condition.

Language assessments

Data from the two language measures (Peabody Picture Vocabulary Test [PPVT] and Test of Language Development [TOLD] Grammatical Understanding subtest) were analyzed using repeated measures linear spline models. There were no statistically detectable differences on these measures for the fall assessment.

There were no statistically detectable differences on either of these measures for the spring pre-kindergarten or spring kindergarten assessment.

Based on the analyses of the two language measures, we conclude that *Project Approach* did not have a statistically detectable effect on language development relative to the control condition.

Behavioral outcomes

We conducted a repeated measures analysis for all three pre-kindergarten social behavioral measures (Social Skills Rating System [SSRS] Social Skills scale, SSRS Problem Behaviors scale, and Preschool Learning Behaviors Scale [PLBS]). The covariates were (a) child's age, (b) gender, (c) race/ethnicity, (d) disability status as reported by the parent, and (e) mother's education. There were no statistically detectable differences on these measures for the fall assessment.

For the spring pre-kindergarten assessment, there were no statistically detectable differences on these measures.

We analyzed the data from the kindergarten versions of the three behavioral measures (SSRS Social Skills scale, SSRS Problem Behaviors scale, and Learning Behaviors Scale [LBS]) using analysis of covariance (ANCOVA). For the ANCOVA analyses, the covariates included (a) the fall pre-kindergarten score of the pre-kindergarten version of the relevant test, along with (b) child's age, (c) gender, (d) race/ethnicity, (e) disability status as reported by the parent, and (f) mother's education.

For the spring kindergarten assessments, teachers rated children who had received *Project Approach* in pre-kindergarten as exhibiting more Problem Behaviors ($ES_s = .49, p < .05$), having weaker Social Skills ($ES_s = -.44, p < .05$), and fewer learning behaviors ($ES_s = -.42, p < .05$), relative to children from the pre-kindergarten control classrooms.

Based on the analyses of the behavioral measures, we conclude that *Project Approach* had a negative effect on children's social and learning behaviors in kindergarten, but not during pre-kindergarten.

Project Approach—Classroom Outcomes

The unadjusted mean scores for classroom measures are reported in table C-13b in appendix C. Covariate adjusted mean differences and standard errors are reported in table D-13b in appendix D. For all analyses of classroom measures, the following variables were included in the model as covariates: (a) teacher has a BA degree, (b) previous teaching experience, (c) child/adult ratio in classroom, (d) average class size, (e) city size, and (f) geographic site. The classroom-level effect sizes (ES_c) are presented in table 11.4.

Overall classroom environment

We conducted a repeated measures analysis on the Early Childhood Environment Rating Scale-Revised (ECERS-R). There was no statistically detectable difference between groups on the fall observation.

No statistically detectable difference between groups was obtained for the spring pre-kindergarten observation.

Based on the analysis of the ECERS-R, we conclude that *Project Approach* did not have a statistically detectable effect on overall classroom quality relative to the control condition.

Teacher-child relationships

We obtained observations on the Arnett Detachment, Harshness, Permissiveness, and Positive Interactions scales in fall and spring of the pre-kindergarten year, and conducted repeated measures analyses. There were no statistically detectable differences on these measures for the fall observation.

There were no statistically detectable differences between groups on the Arnett Detachment, Harshness, Permissiveness, or Positive Interaction scales in spring of the pre-kindergarten year.

Based on the analyses of the four Arnett scales, we conclude that *Project Approach* did not have a statistically detectable effect on teacher-child relationships relative to the control condition.

Classroom instruction

We obtained observations on classroom instruction in (a) early literacy (Teacher Behavior Rating Scale [TBRS] Print and Letter Knowledge and Written Expression scales), (b) phonological awareness (TBRS Phonological Awareness scale), (c) language (TBRS Book Reading and Oral Language scales), and (d) early mathematics (TBRS Math Concepts scale) for the spring pre-kindergarten assessment only. To analyze these data, ANCOVAs were conducted; the covariates were: (a) teacher has a BA degree, (b) previous teaching experience, (c) child/adult ratio in classroom, (d) average class size, (e) city size, and (f) geographic site.

There were no statistically detectable differences between groups on any of the TBRS scales.

Based on the analyses of the TBRS Print and Letter Knowledge and Written Expression scales, we conclude that *Project Approach* did not have a statistically detectable effect on early literacy instruction relative to the control condition.

Based on the analysis of the TBRS Phonological Awareness scale, we conclude that *Project Approach* did not have a statistically detectable effect on instruction in phonological awareness relative to the control condition.

Based on the analysis of the TBRS Book Reading and Oral Language scales, we conclude that *Project Approach* did not have a statistically detectable effect on language instruction relative to the control condition.

Based on the analysis of the TBRS Math Concepts scale, we conclude that *Project Approach* did not have a statistically detectable effect on early mathematics instruction relative to the control condition.

Summary of Findings for *Project Approach*

The findings for *Project Approach* are summarized in table 11.4.

Table 11.4. Effect sizes for *Project Approach*

Measure	Student-level effect sizes (ES _s)		
	RM analysis Spring Pre-K	RM analysis Spring K	ANCOVA Spring K
Mathematics			
WJ Applied Problems	.07	.27	—
CMA-A Mathematics Composite	.18	.22	—
Shape Composition ¹	.27	.24	—
Reading			
TERA	.14	.29	—
WJ Letter Word Identification	.42	.03	—
WJ Spelling	.27	.14	—
Phonological awareness			
Pre-CTOPPP/CTOPP	.05	†	-.17
Language			
PPVT	.16	.10	—
TOLD	.15	.32	—
Behavior			
SSRS Social Skills	.04	†	-.44*
SSRS Problem Behavior ²	.50	†	.49*
PLBS/LBS	-.31	†	-.42*
	Classroom-level effect sizes (ES _c)		
Measure	RM analysis Spring Pre-K	ANCOVA Spring Pre-K	
Global classroom quality			
ECERS-R	-.19	—	
Teacher-child interaction			
Arnett Detachment ³	.57	—	
Arnett Harshness ³	.86	—	
Arnett Permissiveness ³	-.43	—	
Arnett Positive Interactions	-.99	—	
Teacher instructional practices ⁴			
TBRs Book Reading	†	-.76	
TBRs Oral Language	†	-.42	
TBRs Phonological Awareness	†	-1.19	
TBRs Print and Letter Knowledge	†	.34	
TBRs Written Expression	†	.62	
TBRs Math Concepts	†	-.64	

— Not available.

† Not applicable. Four of the kindergarten student-level measures were not on the same scale as the pre-kindergarten measures. The classroom-level data were only collected during the pre-kindergarten year of the study.

* $p < .05$ ¹ Building Blocks, Shape Composition task² Higher scores on this scale represent more negative child behaviors.³ Lower scores on this scale represent a more positive classroom environment.⁴ ANCOVA models for the TBRs measures did not include baseline pretest scores because TBRs data were only collected in spring of the pre-kindergarten year.

NOTE: RM: Repeated Measures

ANCOVA: Analysis of covariance

Significance indications (p -values) in the table refer to the tests of contrasts between intervention and control groups that underlie the effect sizes reported here. Refer to the glossary for abbreviations of the measures.

SOURCE: The Preschool Curriculum Evaluation Research (PCER) Study.

Page intentionally left blank.

Chapter 12. *Project Construct*: University of Missouri-Columbia (Missouri site)

Curriculum

The University of Missouri (Missouri) research team evaluated the *Project Construct* curriculum. *Project Construct* was developed under the direction of the Missouri Department of Elementary and Secondary Education in 1986 to fulfill the need for a curriculum and assessment framework that supports children’s learning. *Project Construct* is derived from constructivism—the theoretical view that learners construct knowledge through interactions with the physical and social environments. The preschool curriculum, the *Early Childhood Framework*, was first published in 1992 by the Project Construct National Center. The *Project Construct* approach is organized around 29 goals for students that are set within a context of four developmental domains:

- Cognitive;
- Representational;
- Sociomoral; and
- Physical.

The Project Construct National Center supports professional development through institutes, workshops, conferences, and on-site consultations as well as through extensive print and video materials.

Sample

The Missouri research team recruited full-day child-care centers through initial phone contacts followed by a letter to briefly explain the study. The program directors were asked to complete a “preschool information form” to clarify enrollment and demographics of the children and staff. If the data on the preschool information form appeared to meet the criteria for eligibility, the director was again contacted. Letters explaining the study and a cooperation agreement were sent to each director and teacher. The primary incentive was free training in *Project Construct* for the treatment group teachers in the initial year of the study and for the control teachers the following year. The treatment classrooms also received supplies and materials to support the implementation of *Project Construct*.

All of the preschools are full-day programs. The preschool program staff assisted with the recruitment of parents and children for the study. The average parental consent rate was 90 percent (90% for the treatment group, 89% for the control group). A total of 231 children and parents were recruited. Data were collected on a total sample of 228 children and 212 parents at the time of the fall baseline data collection.

In the follow-up year of the evaluation, the sample of schools went from 21 in pre-kindergarten to 124 schools in kindergarten. The sample of classrooms went from 23 preschool to 166 kindergarten classrooms. Data were collected on 188 children and 195 parents from the original sample.

Children and Families

The children were 4.7 years old at the time of baseline data collection and less than half (45%) were male. The majority of the sample of preschoolers was White (65%) or African American (25%). Table 12.1 provides additional information on the demographic characteristics of the children in the study sample. There were no statistically detectable differences between the treatment and control groups on the child characteristics.

Table 12.1. Child demographic characteristics for *Project Construct*

Characteristics	Full sample n = 231	Curriculum comparison	
		Control n = 108	Treatment n = 123
Age at baseline (years), mean	4.7	4.7	4.6
Gender (% male)	45.2	45.4	45.0
Race/ethnicity (%)			
White, non-Hispanic	64.8	66.0	63.7
African American, non-Hispanic	25.5	24.3	26.5
Hispanic	2.8	5.8	0.0
Asian or Pacific Islander	‡	‡	0.0
Native American	‡	‡	‡
Multiple/other	5.6	‡	8.0
Child disability status (parent reported, %)	15.1	13.0	17.0

‡ Reporting standards not met. Values suppressed to protect participant confidentiality.

SOURCE: PCER Parent Interview (Fall 2003, Spring 2004, and Spring 2005).

The demographic characteristics of the primary caregivers, who were most often the biological or adoptive mother, are presented in table 12.2. The average age of the primary caregiver was 32 years. About half (52%) were married, and 27 percent were never married. More than half of the primary caregivers reported having some college (36%) or a bachelor's or higher (28%); 27 percent had a high school diploma or GED; and 8 percent had not finished high school. Most (74%) of the primary caregivers were employed full-time; 12 percent were employed part-time; and 12 percent were unemployed. At baseline, mothers in the treatment group were older relative to those assigned to the control group (33 years vs. 31 years, $p < .05$).

Teachers

There were 23 teachers who participated in the preschool year intervention study. All of the teachers were female, and most were White (70%) or African American (26%). On average, the preschool teachers had 10 years of teaching experience, with an average of 8 years of experience teaching preschool. The majority had no college education (61%) and 26 percent had a bachelor's degree. The majority (78%) reported having no teaching credential. Table 12.3 provides additional information on the characteristics of the preschool sample of teachers. There were no statistically detectable differences between the treatment and control groups on the teacher characteristics.

Programs/Classrooms

The average preschool class size was 10.9 children. The child-staff ratio was on average 6.4 children to one teacher or program staff person.

Random Assignment

The Missouri research team identified and recruited a convenience sample of preschools from urban and rural locations in Missouri. Along with the Missouri researchers, Mathematica Policy Research, Inc. (MPR) determined the unit of random assignment for this research site. The MPR research staff randomly assigned preschool centers to treatment and control conditions because a preschool operated only one classroom or it was not feasible to vary the curriculum condition within a school. To increase the precision with which to estimate impacts, MPR grouped schools into blocks of two, and randomly assigned half the schools in each

Table 12.2. Primary caregiver demographic characteristics for *Project Construct*

Characteristics	Full sample n = 212	Curriculum comparison	
		Control n = 100	Treatment n = 112
Age at baseline (years), mean	32.2	31.0	33.3*
Marital status (%)			
Married	51.9	53.0	50.9
Separated/Divorced	19.8	16.0	23.2
Widowed	‡	‡	0.0
Never Married	27.4	29.0	25.9
Race/ethnicity (%)			
White, non-Hispanic	71.7	72.0	71.4
African American, non-Hispanic	25.0	23.0	26.8
Hispanic	2.4	4.0	‡
Asian or Pacific Islander	‡	‡	0.0
Native American	‡	0.0	‡
Multiple/other	0.0	0.0	0.0
Educational level (%)			
Did not finish high school	8.0	11.0	5.4
High school diploma or GED	27.4	25.0	29.5
Some college	36.3	42.0	31.3
College graduate	28.3	22.0	33.9
Employment (%)			
Full-time	74.1	73.0	75.0
Part-time	12.3	9.0	15.2
Unemployed	12.3	17.0	8.0
Other	‡	‡	‡

‡ Reporting standards not met. Values suppressed to protect participant confidentiality.

* $p < .05$

SOURCE: PCER Parent Interview (Fall 2003, Spring 2004, and Spring 2005).

block to the treatment group and half to the control group. The MPR research staff formed blocks by matching schools on easily measured characteristics such as teachers' experience, school location, or score on a state report card system and, in doing so, increased the probability that those characteristics would be evenly distributed between the overall treatment and control groups. MPR staff used a random number function (RAND function in MS Excel) to generate random numbers. They sorted preschools by block and assigned a random number to each preschool. The preschools were then randomly assigned to treatment and control conditions. The staff assigned the highest ranking preschool within the block to the treatment condition, the next highest to the control condition, alternating assignment to treatment and control conditions until all preschools were randomly assigned to one of two conditions. Twenty-three preschool programs (26 preschool classrooms) were initially recruited and randomly assigned to treatment and control conditions. The final study sample of preschool programs included a total of 21 preschool centers (10 control and 11 treatment). The final sample of preschool centers included a sample of 23 preschool classrooms and teachers. There were a total of 11 control classrooms (one preschool center with two classrooms); and 12 treatment classrooms (one preschool center with two classrooms).

Table 12.3. Preschool teacher characteristics for *Project Construct*

Characteristics	Full sample n = 23	Curriculum comparison	
		Control n = 11	Treatment n = 12
Gender (% female)	100.0	100.0	100.0
Race/ethnicity (%)			
White, non-Hispanic	70.0	82.0	58.0
African American, non-Hispanic	26.0	‡	42.0
Hispanic	0.0	0.0	0.0
Asian or Pacific Islander	0.0	0.0	0.0
Native American	0.0	0.0	0.0
Multiple/other	‡	‡	0.0
Educational level (%)			
High school diploma or GED	61.0	55.0	67.0
Associate's degree	‡	0.0	‡
Bachelor's degree	26.0	36.0	‡
Graduate degree	0.0	0.0	0.0
Current teaching license/certificate (%)	‡	‡	‡
Child Development Associate (CDA) (%)	‡	‡	‡
State-awarded preschool certificate (%)	‡	‡	‡
No credential (%)	78.0	82.0	75.0
Years of teaching experience, overall (mean)	9.8	8.1	11.4
Years of preschool teaching experience (mean)	7.6	6.3	8.6

‡ Reporting standards not met. Values suppressed to protect participant confidentiality.

SOURCE: PCER Preschool Teacher Survey (Fall 2003 and Spring 2004).

Contamination

Because schools were assigned to either the treatment or the control group, there was little risk of contamination across the treatment and control conditions.

Control Condition

In the control schools, teacher-developed, generic curricula were implemented.

Data Collection

MPR collected the child, parent, teacher, and school data for the Missouri site for all three waves of data collection. The fall assessment data collection window for child assessments ranged from September 26, 2003 to November 11, 2003. The average delay from the beginning of the treatment (i.e., start of the school year) to the beginning of the fall assessment window was 42 days. The spring pre-kindergarten window was April 5, 2004 to June 20, 2004, and the kindergarten follow-up window was April 18, 2005 to June 8, 2005.

Attrition

A total of 26 classrooms/teachers (13 control and 13 treatment classrooms) were recruited at the beginning of the study. The final sample included 23 teachers and classrooms (11 control and 12 treatment classrooms) because two preschool programs (housing a total of three preschool classrooms) were dropped from the final study sample. One program (two classrooms) was closed and another program (one classroom) was folded into an existing preschool program because of low enrollment numbers. These changes resulted in a loss of

two preschool programs and three preschool classrooms (two control classrooms and one treatment classroom) in fall of the pre-kindergarten year.

For the child assessment, the fall 2003 response rate was 99 percent, the spring 2004 pre-kindergarten response rate was 90 percent, and the kindergarten follow-up response rate was 81 percent.

Implementation

The teachers who were assigned to the *Project Construct* treatment condition received training on three curriculum modules. The training consisted of three 12-hour modules; four 4-hour on-site consultations (the first three shortly after the completion of the modules) in the participants' classrooms; and two 3-hour follow-up workshops (attendance was not mandatory). The modules were designed for educators of children ages 3-5 years. The training on Module 1 (the young child and the learning environment) was held in August 2003; training on Module 2 (early literacy and the expressive arts) was held in October 2003; and training on Module 3 (young children's mathematical and scientific thinking) was held in November 2003. The three modules cover the entire *Project Construct* early childhood framework. The *Project Construct* training institute presented content that is aligned with the Missouri Pre-kindergarten Standards and covers levels 1-3 of the Missouri Core Teacher Competencies.

The onsite consultations occurred following the completion of each module training session. During the onsite consultations, the *Project Construct* consultant observed the teacher in her classroom during a regular classroom session. The consultant then provided the teacher with feedback based on topics from the previous module training and addressed how the training material could be incorporated into the teacher's practice. This feedback included changes in the classroom environment, curriculum planning, family involvement, and teacher-child interaction. Teachers discussed their strengths, areas of needs, goals, and questions or concerns with the consultants.

The *Project Construct* training institute offered follow-up workshops two times during the school year. The subject for each follow-up workshop was a topic identified by the module participants. The workshops provided opportunities for the participants to have interactive and in-depth experiences that provided learning strategies related to the identified topic as well as the opportunity to deepen their understanding of constructivism.

The Missouri research team collected site-specific curriculum fidelity data using the Project Construct Early Childhood Classroom Survey (PC-ECCOS). Observations were conducted in treatment and control classrooms in fall 2003, and spring 2004. The initial curriculum fidelity observation occurred September 19, 2003 to October 29, 2003. The second fidelity check occurred in April and May of 2004. The PC-ECCOS uses a three-point scale to measure evidence for curriculum implementation for each item (1 = *no evidence*; 2 = *some evidence*; 3 = *extensive evidence*). *No evidence* indicates that raters observed *no* evidence of constructivist activities/practices, which would suggest that *Project Construct* was not being implemented in a classroom. *Some evidence* indicates that the raters observed a *fair* amount of evidence for a constructivist approach and teachers are implementing *Project Construct* to some extent. *Extensive evidence* indicates that raters observed a classroom that is *exemplary* in its implementation of constructivism and *Project Construct*.

Implementation Fidelity Ratings

Each research team used a global fidelity measure to rate the overall fidelity with which the curricula were implemented in the preschool year of the project. A four-point scale ranging from "Not at All" (0) to "High" (3) was used to rate each treatment classroom. Researchers were asked to use their site-specific implementation and fidelity data to rate each treatment classroom on the global fidelity measure as High, Medium, Low, or Not at All. Researchers were also asked to provide a global rating for the control group curriculum. The *Project Construct* curriculum was rated at the low Medium level (1.7) on the global

implementation fidelity measure. The control group curriculum was rated at the Medium level (2.3) on the global fidelity measure.

Impact Analysis Results

We begin with the analyses of the child-level measures (i.e., the mathematics, reading, phonological awareness, and language assessments) followed by the analyses of the classroom observation data.

Project Construct—Child Outcomes

The unadjusted mean scores for child-level measures are reported in table C-14a in appendix C. Covariate adjusted mean differences and standard errors are reported in table D-14a in appendix D. For all analyses of child-level measures, the following covariates were included: (a) child's age, (b) gender, (c) race/ethnicity, (d) disability status as reported by parent, and (e) mother's education. The student-level effect sizes (ESs) are presented in table 12.4.

Mathematics assessments

We used repeated measures linear spline models to analyze the data from all three mathematics measures (Woodcock Johnson [WJ] Applied Problems, Child Math Assessment-Abbreviated [CMA-A] Composite Score, and Shape Composition). There were no statistically detectable differences for the fall pre-kindergarten assessment.

There was no statistically detectable difference for the spring pre-kindergarten or kindergarten assessments on the WJ Applied Problems and the CMA-A Composite Score.

There was a statistically reliable negative effect on the Shape Composition scale ($ES = -.42, p < .05$) for spring pre-kindergarten assessments, such that students in the *Project Construct* classrooms had lower overall scores relative to students in the control classrooms. There was no statistically detectable difference in spring of the following year.

Based on the analyses for the three mathematics measures, we conclude that *Project Construct* did not have a statistically detectable effect on mathematics relative to the control condition.

Reading assessments

Data from the three reading measures (Test of Early Reading Ability [TERA], WJ Letter Word Identification, and WJ Spelling) were analyzed using repeated measures linear spline models. There were no statistically detectable differences for the fall pre-kindergarten assessment.

There were no statistically detectable differences on any of these measures for the spring pre-kindergarten or spring kindergarten assessments.

Based on the analyses for the three reading measures, we conclude that *Project Construct* did not have a statistically detectable effect on reading relative to the control condition.

Phonological awareness

The phonological awareness measures were the Preschool Comprehensive Phonological and Print Processing (Pre-CTOPPP), Elision subtest, and the Comprehensive Test of Phonological Processing (CTOPP), Kindergarten, Elision subtest. We conducted a repeated measures analysis on the Pre-CTOPPP fall and spring pre-kindergarten data. There was no statistically detectable difference on the Pre-CTOPPP for the fall assessment.

There was no statistically detectable difference on the Pre-CTOPPP for the spring pre-kindergarten assessment.

We analyzed the kindergarten CTOPP data using analysis of covariance (ANCOVA). For the ANCOVA analysis, the covariates were the (a) Pre-CTOPPP fall assessment score, (b) child's gender, (c) age, (d)

race/ethnicity, (e) disability status as reported by parent, and (f) mother's education. There was no statistically detectable difference between groups on the CTOPP for the spring kindergarten assessment.

Based on the analyses of the Pre-CTOPPP and CTOPP, we conclude that *Project Construct* did not have a statistically detectable effect on phonological awareness relative to the control condition.

Language assessments

Data from the two language measures (Peabody Picture Vocabulary Test [PPVT] and Test of Language Development [TOLD] Grammatical Understanding subtest) were analyzed using repeated measures linear spline models. There were no statistically detectable differences for the fall pre-kindergarten assessment.

There were no statistically detectable differences on any of these measures for the spring pre-kindergarten or spring kindergarten assessments.

Based on the analyses for the two language measures, we conclude that *Project Construct* did not have a statistically detectable effect on language development relative to the control condition.

Behavioral outcomes

We conducted a repeated measures analysis for all three pre-kindergarten social behavioral measures (Social Skills Rating System [SSRS] Social Skills scale, SSRS Problem Behaviors scale, and Preschool Learning Behaviors Scale [PLBS]). The covariates were (a) child's age, (b) gender, (c) disability status as reported by the parent, (d) race/ethnicity, and (e) mother's education. There were no statistically detectable differences on these measures for the fall assessment.

For the spring pre-kindergarten assessment, there were no statistically detectable differences on any of these measures.

We analyzed the data from the kindergarten versions of the three behavioral measures (SSRS Social Skills scale, SSRS Problem Behaviors scale, Learning Behaviors Scale [LBS]) using analysis of covariance (ANCOVA). For the ANCOVA analyses, the covariates included (a) the fall pre-kindergarten score of the pre-kindergarten version of the relevant test, along with (b) child's age, (c) gender, (d) race/ethnicity, (e) disability status as reported by the parent, and (f) mother's education.

There were no statistically detectable differences between groups on any of these measures for the spring kindergarten assessment.

Based on the analyses of the three behavioral measures, we conclude that *Project Construct* did not have a statistically detectable effect on children's social and learning behaviors relative to the control condition.

Project Construct—Classroom Outcomes

The unadjusted mean scores for classroom measures are reported in table C-14b in appendix C. Covariate adjusted mean differences and standard errors are reported in table D-14b in appendix D. For all analyses of classroom measures, the following variables were included in the model as covariates: (a) teacher has a BA degree, (b) previous teaching experience, (c) child/adult ratio in classroom, (d) average class size, (e) city size, and (f) geographic site. The classroom-level effect sizes (ES_c) are presented in table 12.4.

Overall classroom environment

We conducted a repeated measures analysis on the Early Childhood Environment Rating Scale-Revised (ECERS-R). There was no statistically detectable difference between groups on the fall observation.

No statistically detectable difference between groups was obtained for the spring pre-kindergarten observation.

Based on the analysis of the ECERS-R, we conclude that *Project Construct* did not have a statistically detectable effect on overall classroom quality relative to the control condition.

Teacher-child relationships

We obtained observations on the Arnett Detachment, Harshness, Permissiveness, and Positive Interactions scales in fall and spring of the pre-kindergarten year, and conducted repeated measures analyses. There were no statistically detectable differences on these measures for the fall observation.

There were no statistically significant differences between groups on the Arnett Detachment, Harshness, Permissiveness, or Positive Interaction scales for the spring pre-kindergarten observation.

Based on the analyses of the four Arnett scales, we conclude that *Project Construct* did not have a statistically detectable effect on teacher-child relationships relative to the control condition.

Classroom instruction

We obtained observations on classroom instruction in (a) early literacy (Teacher Behavior Rating Scale [TBRS] Print and Letter Knowledge and Written Expression scales); (b) phonological awareness (TBRS Phonological Awareness scale); (c) language (TBRS Book Reading and Oral Language scales); and (d) early mathematics (TBRS Math Concepts scale) for the spring pre-kindergarten observation only. To analyze these data, ANCOVAs were conducted; the covariates were: (a) teacher has a BA degree, (b) previous teaching experience, (c) child/adult ratio in classroom, (d) average class size, (e) city size, and (f) geographic site.

There were no statistically detectable differences between groups on any of the TBRS scales.

Based on the analyses of the TBRS Print and Letter Knowledge and Written Expression scales, we conclude that *Project Construct* did not have a statistically detectable effect on early literacy instruction relative to the control condition.

Based on the analysis of the TBRS Phonological Awareness scale, we conclude that *Project Construct* did not have a statistically detectable effect on instruction in phonological awareness relative to the control condition.

Based on the analysis of the TBRS Book Reading and Oral Language scales, we conclude that *Project Construct* did not have a statistically detectable effect on language instruction relative to the control condition.

Based on the analysis of the TBRS Math Concepts scale, we conclude that *Project Construct* did not have a statistically detectable effect on early mathematics instruction relative to the control condition.

Summary of Findings for *Project Construct*

The findings for *Project Construct* are summarized in table 12.4.

Table 12.4. Effect sizes for *Project Construct*

Measure	Student-level effect sizes (ES _s)		
	RM analysis Spring Pre-K	RM analysis Spring K	ANCOVA Spring K
Mathematics			
WJ Applied Problems	.06	.08	—
CMA-A Mathematics Composite	-.11	-.06	—
Shape Composition ¹	-.42**	.12	—
Reading			
TERA	.00	-.03	—
WJ Letter Word Identification	-.05	.16	—
WJ Spelling	-.15	.00	—
Phonological awareness			
Pre-CTOPPP/CTOPP	.10	†	-.12
Language			
PPVT	.03	.10	—
TOLD	-.05	.01	—
Behavior			
SSRS Social Skills	.22	†	.12
SSRS Problem Behavior ²	-.08	†	.07
PLBS/LBS	.00	†	-.02
Measure	Classroom-level effect sizes (ES _c)		
	RM analysis Spring Pre-K	ANCOVA Spring Pre-K	
Global classroom quality			
ECERS-R	.54	—	
Teacher-child interaction			
Arnett Detachment ³	.12	—	
Arnett Harshness ³	-.13	—	
Arnett Permissiveness ³	-.02	—	
Arnett Positive Interactions	.46	—	
Teacher instructional practices ⁴			
TBRS Book Reading	†	.81	
TBRS Oral Language	†	.52	
TBRS Phonological Awareness	†	.01	
TBRS Print and Letter Knowledge	†	.34	
TBRS Written Expression	†	.43	
TBRS Math Concepts	†	.53	

— Not available.

† Not applicable. Four of the kindergarten student-level measures were not on the same scale as the pre-kindergarten measures. The classroom-level data were only collected during the pre-kindergarten year of the study.

** $p < .01$ ¹ Building Blocks, Shape Composition task² Higher scores on this scale represent more negative child behaviors.³ Lower scores on this scale represent a more positive classroom environment.⁴ ANCOVA models for the TBRS measures did not include baseline pretest scores because TBRS data were only collected in spring of the pre-kindergarten year.

NOTE: RM: Repeated Measures

ANCOVA: Analysis of covariance

Significance indications (p -values) in the table refer to the tests of contrasts between intervention and control groups that underlie the effect sizes reported here. Refer to the glossary for abbreviations of the measures.

SOURCE: The Preschool Curriculum Evaluation Research (PCER) Study.

Page intentionally left blank.

Chapter 13. *Ready, Set, Leap!*: University of California, Berkeley (New Jersey site)

Curriculum

The University of California, Berkeley (California) researchers, in collaboration with RMC Research (RMC), implemented the *Ready, Set, Leap!* curriculum. *Ready, Set, Leap!* is a comprehensive, pre-kindergarten curriculum that combines research-based instructional approaches with multisensory technology. The curriculum is structured around 9 thematic units, each with 120 detailed lesson plans for large- and small-group instruction, and ongoing informal and formal assessment tools. The curriculum aligns with the goals and research requirements of the National Association for the Education of Young Children (NAEYC), the National Head Start Association, and the Early Reading First initiative. This balanced program stresses the importance of active and experiential learning, social and emotional development, teacher-child relationships, and the home-school connection.

All elements are incorporated into the curriculum to provide teachers with comprehensive pre-kindergarten instruction.

The curriculum topics include language and early literacy, mathematics, science, social studies, fine arts, health and safety, personal and social development, physical development, and technology applications.

The curriculum emphasizes the following elements:

- literacy and language development, focusing particularly on scaffolding;
- phonological awareness;
- alphabetic knowledge;
- print awareness;
- oral language development;
- reading aloud; and
- reading comprehension through story discussion.

The technology is designed within each thematic unit to provide center-based activities to integrate the senses of touch, sight, and sound by encouraging students to actively engage with literacy and language, and allowing students to have individualized feedback and support throughout the learning process. There is also a home component to encourage parent-child interactions that support children's learning activities in the preschool setting. The *Ready, Set, Leap!* program application includes family letters, take-home books, and specific strategies specifically for forging strong home-school connections.

Sample

The California research team recruited pre-kindergarten programs in New Jersey. Members of the research team attended a regional pre-kindergarten center meeting with the Director of Early Childhood programs in one large urban area and asked directors to contact them if they were interested in participating in the project. The research team then contacted individual centers that had NAEYC certification. All of the centers offered full-day academic pre-kindergarten programs, typically from about 9:00 a.m. to 3:30 p.m. The exact length of day varied because many of the centers have wrap-around services and children may arrive early in the morning and stay until early evening. From the pool of eligible centers, a total of 39 classrooms/ teachers

were recruited. Treatment and control group teachers and teaching assistants received an incentive for participating in the study.

The research team did not contact families directly but worked with the center directors and classroom teachers to recruit participants. The local site coordinators (i.e., members of the research team) worked with teachers and the director to obtain parental consent for parent and child participation in the study. Teachers asked the parents for their consent to participate in the study. The research team obtained informed consent for 470 parents and children. Approximately 89 percent of the eligible sample of parents and children agreed to participate in the study. The average parental consent rate was 89 percent (93% for the treatment group, 84% for the control group). A final sample of 286 parents and children (149 treatment, 137 control) were included in the study. Data were collected on a sample of 275 children and 261 parents at the time of the fall baseline data collection out of the final baseline sample of 286 parents and children.

In the follow-up year of the evaluation, participants from 21 preschools were followed into 94 schools in kindergarten. The sample of classrooms went from 39 preschool to 162 kindergarten classrooms. Data were collected on 248 children and 218 parents from the original sample.

Children and Families

The children were 4.5 years of age at the time of baseline data collection and more than half (54%) were male. The majority of the sample of preschoolers were African American (78%) or Hispanic (20%). Table 13.1 provides additional information on the demographic characteristics of the children in the study sample. There were no statistically detectable differences between the treatment and control groups on these child characteristics.

The demographic characteristics of the primary caregivers, who were most often the biological or adoptive mother, are presented in table 13.2. The average age of the primary caregiver was 31 years. Nearly two-thirds (63%) were never married, and about a quarter (26%) were married. Forty-three percent of the primary caregivers reported having a high school diploma or GED; 27 percent had some college; 12 percent had a bachelor's degree or higher; and 19 percent did not finish high school. More than half (53%) of the primary caregivers were employed full-time, 13 percent were employed part-time, and 32 percent were unemployed. There were no significant differences between the treatment and control groups on the primary caregiver characteristics.

Table 13.1. Child demographic characteristics for *Ready, Set, Leap!*

Characteristics	Full sample n = 286	Curriculum comparison	
		Control n = 137	Treatment n = 149
Age at baseline (years), mean	4.5	4.5	4.5
Gender (% male)	54.2	56.8	51.7
Race/ethnicity (%)			
White, non-Hispanic	0.0	0.0	0.0
African American, non-Hispanic	78.4	74.8	81.6
Hispanic	20.1	22.8	17.6
Asian or Pacific Islander	‡	‡	0.0
Native American	‡	‡	‡
Multiple/other	‡	‡	0.0
Child disability status (parent reported, %)	8.1	8.2	8.0

‡ Reporting standards not met. Values suppressed to protect participant confidentiality.

SOURCE: PCER Parent Interview (Fall 2003, Spring 2004, and Spring 2005).

Table 13.2. Primary caregiver demographic characteristics for *Ready, Set, Leap!*

Characteristics	Full sample n = 256	Curriculum comparison	
		Control n = 120	Treatment n = 136
Age at baseline (years), mean	30.8	30.9	30.8
Marital status (%)			
Married	26.2	25.8	26.5
Separated/Divorced	10.2	12.5	8.1
Widowed	‡	0.0	‡
Never Married	63.3	61.7	64.7
Race/ethnicity (%)			
White, non-Hispanic	‡	‡	0.0
African American, non-Hispanic	77.0	76.3	77.6
Hispanic	21.4	21.2	21.6
Asian or Pacific Islander	‡	‡	0.0
Native American	‡	‡	‡
Multiple/other	0.0	0.0	0.0
Educational level (%)			
Did not finish high school	18.6	21.2	16.3
High school diploma or GED	42.7	39.8	45.2
Some college	26.5	27.1	25.9
College graduate	12.3	11.9	12.6
Employment (%)			
Full-time	52.7	55.8	50.0
Part-time	12.9	13.3	12.5
Unemployed	32.4	27.5	36.8
Other	‡	3.3	‡

‡ Reporting standards not met. Values suppressed to protect participant confidentiality.
SOURCE: PCER Parent Interview (Fall 2003, Spring 2004, and Spring 2005).

Teachers

There were 39 teachers who participated in the preschool year intervention study. Most (95%) of the preschool teachers were female. The racial/ethnic composition of the sample included African American (61%), Hispanic (18%), and White (10%) teachers. The preschool teachers had an average of 8 years of teaching experience and 5 years of experience teaching preschool. The majority of teachers had a bachelor's degree (69%); 10 percent had an associate's degree; and 15 percent had no college degree. The majority of teachers reported having a current teaching license/certificate (51%), a state-awarded preschool certificate (49%), or a Child Development Associate (CDA) credential (26%). Table 13.3 provides additional information on the characteristics of the preschool sample of teachers. There were no statistically detectable differences between the treatment and control groups on the teacher characteristics.

Programs/Classrooms

The average preschool class size was 12.3 children. The child-staff ratio was an average of six children to one teacher or program staff person.

Table 13.3. Preschool teacher characteristics for *Ready, Set, Leap!*

Characteristics	Full sample n = 39	Curriculum comparison	
		Control n = 18	Treatment n = 21
Gender (% female)	95.0	94.0	95.0
Race/ethnicity (%)			
White, non-Hispanic	11.0	12.0	‡
African American, non-Hispanic	61.0	53.0	67.0
Hispanic	18.0	‡	19.0
Asian or Pacific Islander	‡	0.0	‡
Native American	0.0	0.0	0.0
Multiple/other	‡	‡	0.0
Educational level (%)			
High school diploma or GED	15.0	11.0	19.0
Associate's degree	10.0	17.0	‡
Bachelor's degree	69.0	72.0	67.0
Graduate degree	‡	0.0	‡
Current teaching license/certificate (%)	51.0	39.0	62.0
Child Development Associate (CDA) (%)	26.0	22.0	29.0
State-awarded preschool certificate (%)	49.0	39.0	57.0
No credential (%)	18.0	28.0	‡
Years of teaching experience, overall (mean)	8.0	6.8	9.0
Years teaching preschool (mean)	5.4	4.4	6.3

‡ Reporting standards not met. Values suppressed to protect participant confidentiality.

SOURCE: PCER Preschool Teacher Survey (Fall 2003 and Spring 2004).

Random Assignment

The California research team identified and recruited a convenience sample of 21 preschools from an urban area in New Jersey. Along with the California research team, Mathematica Policy Research, Inc. (MPR) determined the unit of random assignment at the research study site in New Jersey. The MPR research staff randomly assigned individual classrooms to conditions after it was determined that the experimental curriculum could be introduced in one classroom without affecting neighboring classrooms in the same school, and that preschool staff were willing to use different curricula within the same setting. To increase the precision with which to estimate impacts, MPR grouped classrooms into blocks of two or more and randomly assigned half the classrooms in each block to the treatment group and half to the control group. MPR research staff formed blocks by matching schools on easily measured characteristics such as teachers' experience, school location, or score on a state report card system and, in doing so, increased the probability that those characteristics would be evenly distributed between the overall treatment and control groups. MPR staff used a random number function (RAND function in MS Excel) to generate random numbers. They sorted the classrooms by block and assigned a random number to each classroom. The classrooms were then randomly assigned to treatment and control conditions. The staff assigned the highest ranking classroom within the block to the treatment condition, the next highest to the control condition, alternating assignment to treatment and control conditions until all classrooms were randomly assigned to one of two conditions. A total of 39 classrooms (21 treatment and 18 control) were randomly assigned to conditions. The 39 classrooms were drawn from 21 schools (10 schools and 18 control classrooms; 11 schools and 21 treatment classrooms).

Contamination

Because both *Ready Set Leap!* and control classrooms could reside within the same school, the researchers monitored the classrooms to ensure that treatment group teachers were not sharing materials and instructional practices with the control group teachers.

Control Condition

In the control condition, the teachers used the *High/Scope* approach to early childhood education.

Data Collection

MPR collected the child, parent, teacher, and school data for the New Jersey site for all three waves of data collection. The fall assessment data collection window for child assessments ranged from October 20, 2003 to November 19, 2003. The average delay from the beginning of the treatment (i.e., start of the school year) to the beginning of the fall assessment window was 35 days. The spring pre-kindergarten window was May 10, 2004 to June 15, 2004, and the kindergarten follow-up window was April 25, 2005 to June 8, 2005.

Attrition

Thirty-nine classrooms were randomly assigned to treatment and control conditions. All 39 classrooms remained in the study throughout the pre-kindergarten year.

For the child assessment, the fall 2003 response rate was 96 percent; the spring 2004 pre-kindergarten response rate was 92 percent; and the kindergarten follow-up response rate was 87 percent.

Implementation

The *Ready, Set, Leap!* curriculum was implemented in the 21 treatment classrooms in September 2003. Treatment group teachers received 4 full days of professional development training. The training sessions were scheduled to occur throughout the preschool year (September 2003, November 2003, January 2004, and March 2004). Curriculum fidelity was measured by triangulating three sources of data: (1) coaching visits that occurred three times during the school year; (2) site coordinator ratings based upon their three visits to each treatment and control classroom; and (3) modified CLASSIC observation coding based upon the site coordinator observations that included a 90-second time sampling procedure, with 32 events recorded for each observation, for a total of 117 observations. Inter-rater reliability was established for approximately 10 percent of the observations.

Implementation Fidelity Ratings

Each research team used a global fidelity measure to rate the overall fidelity with which the curricula were implemented in the preschool year of the project. A four-point scale ranging from “Not at All” (0) to “High” (3) was used to rate each treatment classroom. Researchers were asked to use their site-specific implementation and fidelity data to rate each treatment classroom on the global fidelity measure as High, Medium, Low, or Not at All. Researchers were also asked to provide a global rating for the control group curriculum. The *Ready, Set, Leap!* curriculum (1.9) and the control curriculum (2.0) were both rated at the Medium level on the global implementation fidelity measure.

Impact Analysis Results

We begin with the analyses of the child-level measures (i.e., the mathematics, reading, phonological awareness, and language assessments) and then present the analyses of the classroom observation data.

Ready, Set, Leap!—Child Outcomes

The unadjusted mean scores for child-level measures are reported in table C-15a in appendix C. Covariate adjusted mean differences and standard errors are reported in table D-15a in appendix D. For all analyses of child-level measures, the following covariates were included: (a) child's age, (b) gender, (c) race/ethnicity, (d) disability status as reported by parent, and (e) mother's education. The student-level effect sizes (ESs) are presented in table 13.4.

Mathematics assessments

We used repeated measures linear spline models to analyze the data from all three mathematics measures (Woodcock Johnson [WJ] Applied Problems, Child Math Assessment-Abbreviated [CMA-A] Composite Score, and Shape Composition). There were no statistically detectable differences between the treatment and control groups on the WJ Applied Problems for the fall pre-kindergarten, spring pre-kindergarten, and spring kindergarten assessments.

For the CMA-A Composite Score, there was no statistically detectable difference for the fall pre-kindergarten assessment. There was a statistically significant negative effect (ESs = $-.24$, $p < .05$) for the spring pre-kindergarten assessment, indicating that children in the *Ready, Set, Leap!* classrooms were outperformed by students in the control classrooms. There was no statistically detectable difference between the treatment and control groups for the spring kindergarten assessment.

For the Shape Composition scale, there was a statistically reliable difference favoring the *Ready, Set, Leap!* group on the fall assessment (ESs = $.25$, $p < .05$; follow-up analyses for this finding are included in appendix A). There was no statistically detectable difference between the treatment and control groups for the spring pre-kindergarten or spring kindergarten assessments.

Based on the analyses for the three mathematics measures, we conclude that *Ready, Set, Leap!* did not have a statistically detectable effect on mathematics relative to the control condition.

Reading assessments

Data from the three reading measures (Test of Early Reading Ability [TERA], WJ Letter Word Identification, and WJ Spelling) were analyzed using repeated measures linear spline models. There were no statistically detectable differences for the fall assessment on these measures.

There were no statistically detectable differences on any of these measures for the spring pre-kindergarten or spring kindergarten assessments.

Based on the analyses for the three reading measures, we conclude that *Ready, Set, Leap!* did not have a statistically detectable effect on reading relative to the control condition.

Phonological awareness

The phonological awareness measures were the Preschool Comprehensive Phonological and Print Processing (Pre-CTOPPP), Elision subtest, and the Comprehensive Test of Phonological Processing (CTOPP), Kindergarten, Elision subtest. We conducted a repeated measures analysis on the Pre-CTOPPP fall and spring pre-kindergarten data. There was no statistically detectable difference on the Pre-CTOPPP for the fall assessment.

There was no statistically detectable difference on the Pre-CTOPPP for the spring pre-kindergarten assessment.

We analyzed the kindergarten CTOPP data using analysis of covariance (ANCOVA). For the ANCOVA analysis, the covariates were the (a) Pre-CTOPPP fall assessment score, (b) child's gender, (c) age, race/ethnicity, (d) disability status as reported by parent, and (e) mother's education. There was no statistically detectable difference between groups on the CTOPP for the spring kindergarten assessment.

Based on the analyses of the Pre-CTOPPP and CTOPP, we conclude that *Ready, Set, Leap!* did not have a statistically detectable effect on phonological awareness relative to the control condition.

Language assessments

Data from the two language measures (Peabody Picture Vocabulary Test [PPVT] and Test of Language Development [TOLD] and Grammatic Understanding subtest) were analyzed using repeated measures linear spline models. There were no statistically detectable differences on these measures for the fall pre-kindergarten assessment.

In the spring of the pre-kindergarten and kindergarten years, there were no statistically detectable differences between groups on either measure.

Based on the analyses of the two language measures, we conclude that *Ready, Set, Leap!* did not have a statistically detectable effect on language development relative to the control condition.

Behavioral outcomes

We conducted a repeated measures analysis for all three pre-kindergarten social behavioral measures (Social Skills Rating System [SSRS] Social Skills scale, SSRS Problem Behaviors scale, and Preschool Learning Behaviors [PLBS]). The covariates were (a) child's age, (b) gender, and (c) race/ethnicity, (d) disability status as reported by the parent, and (e) mother's education. There were no statistically detectable differences on these measures for the fall assessment.

For the spring pre-kindergarten assessment, there were no statistically detectable differences on any of these measures.

We analyzed the data from the kindergarten versions of the three behavioral measures (SSRS Social Skills scale, SSRS Problem Behaviors scale, and Learning Behaviors Scale [LBS]) using analysis of covariance (ANCOVA). For the ANCOVA analyses, the covariates included (a) the fall pre-kindergarten score of the pre-kindergarten version of the relevant test, along with (b) child's age, (c) gender, (d) race/ethnicity, (e) disability status as reported by the parent, and (f) mother's education.

There were no statistically detectable differences between groups on any of these measures for the spring kindergarten assessment.

Based on the analyses of the three behavioral measures, we conclude that *Ready, Set, Leap!* did not have a statistically detectable effect on children's social and learning behaviors relative to the control condition.

Ready, Set, Leap!—Classroom Outcomes

The unadjusted mean scores for classroom measures are reported in table C-15b in appendix C. Covariate adjusted mean differences and standard errors are reported in table D-15b in appendix D. For all analyses of classroom measures, the following variables were included in the model as covariates: (a) teacher has a BA degree, (b) previous teaching experience, (c) child/adult ratio in classroom, (d) average class size, (e) city size, and (f) geographic site. The classroom-level effect sizes (ES_c) are presented in table 13.4.

Overall classroom environment

We conducted a repeated measures analysis on the Early Childhood Environment Rating Scale-Revised (ECERS-R). There was no statistically significant difference between groups on the fall observation. No statistically detectable difference between groups was obtained for the spring pre-kindergarten observation.

Based on the analysis of the ECERS-R, we conclude that *Ready, Set, Leap!* did not have a statistically detectable effect on overall classroom quality relative to the control condition.

Teacher-child relationships

We obtained observations on the Arnett Detachment, Harshness, Permissiveness, and Positive Interactions scales in fall and spring of the pre-kindergarten year, and conducted repeated measures analyses. There were no statistically detectable differences on these measures for the fall observation.

There were no statistically detectable differences between groups on the Arnett Detachment, Harshness, Permissiveness, or Positive Interaction scales for the spring pre-kindergarten observation.

Based on the analyses of the four Arnett scales, we conclude that *Ready, Set, Leap!* did not have a statistically detectable effect on teacher-child relationships relative to the control condition.

Classroom instruction

We obtained observations on classroom instruction in (a) early literacy (Teacher Behavior Rating Scale [TBRS] Print and Letter Knowledge and Written Expression scales); (b) phonological awareness (TBRS Phonological Awareness scale); (c) language (TBRS Book Reading and Oral Language scales); and (d) early mathematics (TBRS Math Concepts scale) for the spring pre-kindergarten assessment only. To analyze these data, ANCOVAs were conducted; the covariates were: (a) teacher has a BA degree, (b) previous teaching experience, (c) child/adult ratio in classroom, (d) average class size, (e) city size, and (f) geographic site.

There were no statistically detectable differences between groups on any of the TBRS scales.

Based on the analyses of the TBRS Print and Letter Knowledge and Written Expression scales, we conclude that *Ready, Set, Leap!* did not have a statistically detectable effect on early literacy instruction relative to the control condition.

Based on the analysis of the TBRS Phonological Awareness scale, we conclude that *Ready, Set, Leap!* did not have a statistically detectable effect on instruction in phonological awareness relative to the control condition.

Based on the analysis of the TBRS Book Reading and Oral Language scales, we conclude that *Ready, Set, Leap!* did not have a statistically detectable effect on language instruction relative to the control condition.

Based on the analysis of the TBRS Math Concepts scale, we conclude that *Ready, Set, Leap!* did not have a statistically detectable effect on early mathematics instruction relative to the control condition.

Summary of Findings for *Ready, Set, Leap!*

The findings for *Ready, Set, Leap!* are summarized in table 13.4.

Table 13.4. Effect sizes for *Ready, Set, Leap!*

Measure	Student-level effect sizes (ES _s)		
	RM analysis Spring Pre-K	RM analysis Spring K	ANCOVA Spring K
Mathematics			
WJ Applied Problems	.04	.00	—
CMA-A Mathematics Composite	-.24*	-.10	—
Shape Composition ¹	.08	.03	—
Reading			
TERA	.08	.01	—
WJ Letter Word Identification	.01	-.12	—
WJ Spelling	.20	.04	—
Phonological awareness			
Pre-CTOPPP/CTOPP	-.09	†	-.02
Language			
PPVT	.15	-.02	—
TOLD	-.11	-.03	—
Behavior			
SSRS Social Skills	-.05	†	-.03
SSRS Problem Behavior ²	-.03	†	.07
PLBS/LBS	.07	†	-.01
Measure	Classroom-level effect sizes (ES _c)		
	RM analysis Spring Pre-K	ANCOVA Spring Pre-K	
Global classroom quality			
ECERS-R	.16	—	
Teacher-child interaction			
Arnett Detachment ³	.19	—	
Arnett Harshness ³	.30	—	
Arnett Permissiveness ³	-.24	—	
Arnett Positive Interactions	.04	—	
Teacher instructional practices ⁴			
TBRS Book Reading	†	-.18	
TBRS Oral Language	†	-.24	
TBRS Phonological Awareness	†	.22	
TBRS Print and Letter Knowledge	†	-.02	
TBRS Written Expression	†	.10	
TBRS Math Concepts	†	-.10	

— Not available.

† Not applicable. Four of the kindergarten student-level measures were not on the same scale as the pre-kindergarten measures. The classroom-level data were only collected during the pre-kindergarten year of the study.

* $p < .05$

¹ Building Blocks, Shape Composition task

² Higher scores on this scale represent more negative child behaviors.

³ Lower scores on this scale represent a more positive classroom environment.

⁴ ANCOVA models for the TBRS measures did not include baseline pretest scores because TBRS data were only collected in spring of the pre-kindergarten year.

NOTE: RM: Repeated Measures

ANCOVA: Analysis of covariance

Significance indications (p -values) in the table refer to the tests of contrasts between intervention and control groups that underlie the effect sizes reported here. Refer to the glossary for abbreviations of the measures.

SOURCE: The Preschool Curriculum Evaluation Research (PCER) Study.

Page intentionally left blank.

References

- Abbott-Shim, M., and Sibley, A. (2001). *Assessment Profile for Early Childhood Programs: Research Edition II*. Atlanta, GA: Quality Counts, Inc.
- Abrams & Company (2000). *Let's Read with the Letter People*. Waterbury, CT: Abrams & Company.
- Administration for Children and Families (ACF) (2002a). A Descriptive Study of Head Start Families: *FACES Technical Report I*. Washington, DC: U.S. Department of Health and Human Services.
- Administration for Children and Families (ACF) (2002b). Head Start FACES 2000: A Whole-Child Perspective on Program Performance. Washington, DC: U.S. Department of Health and Human Services.
- Arnett, J. (1989). Caregivers in Day-Care Centers: Does Training Matter? *Journal of Applied Developmental Psychology, 10*: 541-552.
- Bradley, R.H., and Caldwell, B. (1984). The HOME Inventory and Family Demographics. *Developmental Psychology, 20*(2): 315-320.
- Bunce, B.H. (1995). *Building a Language-Focused Curriculum for the Preschool Classroom: Volume II*. Baltimore: Paul H. Brookes.
- Campbell, F.A., Ramey, C.T., Pungello, E.P., Sparling, J., and Miller-Johnson, S. (2002). Early Childhood Education: Young Adult Outcomes from the Abecedarian Project. *Applied Developmental Science, 6*: 42-57.
- Chaney, C. (1992). Language Development, Metalinguistic Skills, and Print Awareness in 3-year-old Children. *Applied Psycholinguistics, 13*: 485-514.
- Clements, D.H., and Sarama, J. (2003). DLM Early Childhood Express Math Resource Guide. Columbus, OH: SRA/McGraw-Hill.
- Clements, D., Sarama, J., and Liu, X. (in press). Development of a Measure of Early Mathematics Achievement Using the Rasch Model: The Research-Based Early Math Assessment. *Educational Psychology*.
- Cost, Quality, and Child Outcomes Study Team (1995). *Cost, Quality, and Child Outcomes in Child Care Centers, Technical Report*. University of Colorado at Denver, CO.
- Cunningham, A.E., and Stanovich, K.E. (1997). Early Reading Acquisition and its Relation to Reading Experience and Ability Ten Years Later. *Developmental Psychology, 33*: 934-945.
- Cunningham, A.E., Stanovich, K.E., and West, K.E. (1994). Literacy Environment and the Development of Children's Cognitive Skills. In E.M. Assink (Ed.), *Literacy Acquisition and Social Context* (pp. 70-90). Oxford, England: Harvester Wheatsheaf/Prentice Hall.

References

- Dempster, A.P., Laird, N.M., and Rubin, D.B. (1977). Maximum Likelihood from Incomplete Data via the EM Algorithm. *Journal of the Royal Statistical Society, Series B*, 39: 1-38.
- Dickinson, D.K., and Tabors, P.O. (2001). *Beginning Literacy with Language: Young Children Learning at Home and School*. Baltimore: Brookes Publishing.
- Dodge, D.T., Colker, L.J., and Heroam, C. (2002). *The Creative Curriculum for Preschool* (4th ed.). Washington, DC: Teaching Strategies, Inc.
- Downer, J., and Pianta, R. (2006). Academic and Cognitive Functioning in First Grade: Associations with Earlier Home and Child Care Predictors and with Concurrent Home and Classroom Experiences. *School Psychology Review*, 35: 11-30.
- Duncan, G.J., Dowsett, C.J., Claessens, A., Magnuson, K., Huston, A.C., Klebanov, P., Pagani, L., Feinstein, L., Engel, A., Brooks-Gunn, J., Sexton, H., Duckworth, K., and Japel, C. (2006). School Readiness and Later Achievement. *Developmental Psychology*, 43(6): 1,428-1,446.
- Dunn, L., and Dunn, L. (1997). *Peabody Picture Vocabulary Test, Third Edition (PPVT-III): Form IIA*. Circle Pines, MN: American Guidance Service, Inc.
- Echols, L.D., West, R.F., Stanovich, K.E., and Zehr, K.S. (1996). Using Children's Literacy Activities to Predict Growth in Verbal Cognitive Skills: A Longitudinal Investigation. *Journal of Educational Psychology*, 88: 296-304.
- Ehri, L.C., and Wilce, L.S. (1980). The Influence of Orthography on Readers' Conceptualization of the Phonemic Structure of Words. *Applied Psycholinguistics*, 1: 371-385.
- Entwisle, D.R., and Alexander, K.L. (1988). Factors Affecting Achievement Test Scores and Marks of Black and White First Graders. *The Elementary School Journal*, 88: 449-471.
- Entwisle, D.R., and Alexander, K.L. (1992). Summer Setback: Race, Poverty, School Composition, and Mathematics Achievement in the First Two Years of School. *American Sociological Review*, 57: 72-84.
- Feise, R.J. (2002). Do Multiple Outcome Measures Require *p*-value Adjustment. *BMC Medical Research Methodology*, 2: 8-11.
- Fitzmaurice, G.M., Laird, N.M., and Ware, J.H. (2004). *Applied Longitudinal Analysis*. New York: John Wiley & Sons.
- Gomez, E.V., Schaalje, G.B., and Fellingham, G.W. (2005). Performance of the Kenward-Roger Method When the Covariance Structure is Selected Using AIC and BIC. *Communications in Statistics: Simulation and Computation*, 34: 377-392.
- Gresham, F.M., and Elliott, S.N. (1990). *Social Skills Rating System*. Circle Pines, MN: American Guidance Service, Inc.
- Harms, T., Clifford, R.M., and Cryer, D. (1998). *Early Childhood Environment Rating Scale. Revised Edition*. New York: Teachers College Press.

References

- Hart, B., and Risley, T.R. (1995). *Meaningful Differences in the Everyday Experience of Young American Children*. Baltimore: Paul H. Brookes Publishing Co.
- Harville, D.A. (1974). Bayesian Inference for Variance Components Using Only Error Contrasts. *Biometrika*, 61: 383-385.
- Harville, D.A. (1977). Maximum Likelihood Approaches to Variance Component Estimation and to Related Problems. *Journal of the American Statistical Association*, 72: 320-338.
- Helm, J.H., and Katz, L. (2001). *Young Investigators: The Project Approach in the Early Years*. New York: Teachers College Press.
- Henderson, C. (1984). *Applications of Linear Models in Animal Breeding*. Ontario, Canada: University of Guelph.
- Henderson, C.R. (1990). Statistical methods in animal improvement: historical overview. In D. Gianola and K. Hammond (Eds.), *Advances in statistical methods for genetic improvement of livestock*. Berlin: Springer-Verlag.
- High Reach Learning (1997a). *High Reach Learning Curriculum for Pre-K: Butterflies*. Charlotte, NC: High Reach Learning.
- High Reach Learning (1997b). *High Reach Learning Curriculum for Pre-K: Framework*. Charlotte, NC: High Reach Learning.
- Hohmann, M., and Weikart, D.P. (2002). *Educating Young Children: Active Learning Practices for Preschool and Child Care Programs*. Ypsilanti, MI: High/Scope Press.
- Juel, C. (1988). Learning to Read and Write: A Longitudinal Study of 54 Children from First Through Fourth Grades. *Journal of Educational Psychology*, 78: 243-255.
- Juel, C. (1991). Beginning Reading. In R. Barr, M.L. Kamil, P.B. Mosenthal, and P.D. Pearson (Eds.), *Handbook of Reading Research*, 2: 759-788. Mahwah, NJ: Lawrence Erlbaum Associates.
- Katz, L.G., and Chard, S.C. (1989). *Engaging Children's Minds: The Project Approach*. Greenwich, CT: Ablex.
- Kenward, M.G., and Roger, J.H. (1997). Small Sample Inference for Fixed Effects from Restricted Maximum Likelihood. *Biometrics*, 53: 983-997.
- Keselman, H.J., Algina, J., Kowalchuk, R.K., and Wolfinger, R.D. (1998). A Comparison of Two Approaches for Selecting Covariance Structures in the Analysis of Repeated Measurements, *Communications in Statistics: Simulation and Computation*, 27: 591-604.
- Kirk, R.E. (1995). *Experimental design: Procedures for the behavioral sciences*. Pacific Grove, CA: Brooks/Cole Publishing Company.
- Klein, A., and Starkey, P. (2000). *Child Math Assessment*. Berkeley, CA: Author.
- Klein, A., and Starkey, P. (2002). *Child Math Assessment-Abbreviated*. Berkeley, CA: Author.

References

- Klein, A., Starkey, P., and Ramirez, A.B. (2002). *Pre-K Mathematics Curriculum*. United States of America: Scott Foresman.
- Kreft, I.I., and de Leeuw, J. (1998). *Introducing Multilevel Modeling*. London: Sage Publications.
- Landry, S.H., Crawford, A., Gunnewig, S., and Swank, P. (2002). *Teacher Behavior Rating Scale (TBR)*. Center for Improving the Readiness of Children for Learning and Education, unpublished research instrument.
- LeapFrog School House. (2003) *Ready, Set, Leap!* Emeryville, CA: Author.
- Lentz, F.E. (1988). Effective Reading Interventions in the Regular Classroom. In J.L. Graden, J.E. Zins, and M.J. Curtis (Eds.), *Alternative Educational Delivery Systems: Enhancing Instructional Options for All Students*. Washington, DC: National Association of School Psychologists.
- Lieberman, I.Y., Shankweiler, D., Fischer, F.W., and Carter, B. (1974). Explicit Syllable and Phoneme Segmentation in the Young Child. *Journal of Experimental Child Psychology*, 18: 201-212.
- Littell, R.C., Milliken, G.A., Stroup, W.W., and Wolfinger, R.D. (1996). *SAS System for Mixed Models*, Cary, NC: SAS Institute Inc.
- Lonigan, C.J., and Farver, J.M. (2002) *Literacy Express: A Preschool Emergent Literacy Curriculum*. Tallahassee, FL: Author.
- Lonigan, C.J., Wagner, R.K., Torgesen, J.K., and Rashotte, C.A. (2002). *The Preschool Comprehensive Test of Phonological and Print Processing*. Tallahassee, FL: Florida State University.
- Mariner, C.L., Zaslow, M.J., and Sugland, B.W. (1998). *Factor Structure and Predictive Validity of the HOME-Short Form for Three Racial/Ethnic Groups in the National Longitudinal Survey of Youth-Child Supplement*. Methods Working Paper. Washington, DC: Child Trends, Inc.
- Mason, J.M., and Stewart, J. (1989). *CAP Early Childhood Diagnostic Instrument, (prepublication edition)*. Iowa City, IA: American Testronics.
- McCulloch, C.E., and Searle, S.R. (2001). *Generalized, Linear, and Mixed Models*. New York: John Wiley & Sons.
- McDermott, P.A., Green, L.F., Francis, J.M., and Stott, D.H. (1999). *Learning Behaviors Scale*. Philadelphia: Edumetric and Clinical Science.
- McDermott, P.A., Green, L.F., Francis, J.M., and Stott, D.H. (2000). *Preschool Learning Behaviors Scale*. Philadelphia: Edumetric and Clinical Science.
- McDermott, P.A., Leigh, N.M., and Perry, M.A. (2002). Development and Validation of the Preschool Learning Behaviors Scale. *Psychology in the Schools*, 39(4): 353-36.
- McGrew, K.S., and Woodcock, R.W. (2001). *Woodcock Johnson Achievement Test-3rd Edition (WJ III). Technical Manual*. Itasca, IL: Riverside Publishing.

References

- Miles, S., and Stipek, D. (2006). Contemporaneous and Longitudinal Associations Between Social Behavior and Literacy Achievement in a Sample of Low-income Elementary School Children. *Child Development*, 77: 103-117.
- Missouri Department of Elementary and Secondary Education. (1992). Project Construct: A Framework for Curriculum and Assessment. Jefferson City, MO: Author.
- National Research Council. (2001). *Eager to Learn: Educating Our Preschoolers*. In B.T. Bowman, M.S. Donovan, and M.S. Burns (Eds.), Committee on Early Childhood Pedagogy. *Commission on Behavioral and Social Sciences and Education*. Washington, DC: National Academy Press.
- Natriello, G., McDill, E.L., and Pallas, A.M. (1990). *Schooling Disadvantaged Children: Racing Against Catastrophe*. New York: Teachers College Press.
- Newcomer, P.L., and Hammill, D.D. (1997). *Examiner's Manual. Test of Language Development (TOLD) Primary: Third Edition*. Austin, TX: Pro-Ed.
- Notari-Syverson, A., O'Connor, R.E., and Vadasy, P.F. (1998). *Ladders to Literacy: A Preschool Activity Book*. Baltimore: Paul H. Brookes Publishing Company.
- Patterson, H.D., and Thompson, R. (1971). Recovery of Interblock Information When Block Sizes are Unequal. *Biometrika*, 58: 545-554.
- Peisner-Feinberg, E.S., and Burchinal, M.R. (1997). Relations Between Preschool Children's Child-care Experiences and Concurrent Development: The Cost, Quality, and Outcomes Study. *Merrill-Palmer Quarterly*, 43: 451-477.
- Pellin, B., and Edmonds, E. (2001). *Bright Beginnings, 1997-2001*. Charlotte, NC: Charlotte-Mecklenburg Schools.
- Perfetti, C.A., Beck, I., Bell, L., and Hughes, C. (1987). Phonemic Knowledge and Learning to Read are Reciprocal: A Longitudinal Study of First Grade Children. *Merrill-Palmer Quarterly*, 33: 283-319.
- Perie, M., Grigg, W.S., and Donahue, P.L. (2005). *The Nation's Report Card: Mathematics 2005*. U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1990-2005 Mathematics Assessments.
- Phelps, P.C. (2002). *Beyond Centers & Circle Time Curriculum*. Lewisville, NC: Kaplan Early Learning Company.
- Princiotta, D., Flanagan, K.D., and Germino-Hausken, E. (2006). *Fifth Grade: Findings From The Fifth-Grade Follow-up of the Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K)* (NCES 2006-038) U.S. Department of Education. Washington, DC: National Center for Education Statistics.
- Radloff, L.S. (1977). The CES-D Scale: A Self-Report Depression Scale for Research in the General Population. *Applied Psychological Measurement*, 1: 385-401.
- Reid, D.K., Hresko, W.P., and Hammill, D.D. (2001). *Test of Early Reading Ability-Third Edition*. Circle Pines, MN: American Guidance Service, Inc.

References

- Rice, M.L., and Wilcox, K.A. (1995). *Building a Language-Focused Curriculum for the Preschool Classroom: Volume I*. Baltimore: Paul H. Brookes.
- Rothman, K. (1990). No Adjustments are Needed for Multiple Comparisons. *Epidemiology*, 1: 43-46.
- Ruopp, R., Travers, J., Glantz, F.M., and Coelen, C. (1979). Children at the Center: Summary of Findings and Their Implications. *Final Report for the National Day Care Study, 1*. Cambridge, MA: Abt Associates.
- Savitz, D.A., and Olshan, A.F. (1995). Multiple Comparisons and Related Issues in the Interpretation of Epidemiologic Data. *American Journal of Epidemiology*, 142: 904-8.
- Scarborough, H.S. (1989). Prediction of Reading Disability from Familial and Individual Differences. *Journal of Educational Psychology*, 81(1): 101-108.
- Schaalje, G.B., McBride, J.B., and Fellingham, G.W. (2002). Adequacy of Approximations to Distributions of Test Statistics in Complex Mixed Linear Models. *Journal of Agricultural, Biological, and Environmental Statistics*. 7: 512-524.
- Schochet, P. (2005). *Statistical Power for Random Assignment Evaluations of Education Programs*. Princeton, NJ: Mathematica Policy Research, Inc.
- Schweinhart, L. (2004). *The High/Scope Perry Preschool Study Through Age 40: Summary, Conclusions, and Frequently Asked Questions*. Ypsilanti, MI: High/Scope Press.
- Searle, S.R., Casella, G.S., and McCulloch, C.E. (1992). *Variance Components*. New York: John Wiley & Sons.
- Shankweiler, D., Crain, S., Katz, L., and Fowler, A.E. (1995). Cognitive Profiles of Reading-disabled Children: Comparison of Language Skills in Phonology, Morphology, and Syntax. *Psychological Science*, 6(3): 149-156.
- Snow, C.E. (Ed.). (1998). *Preventing Reading Difficulties in Young Children*. Washington, DC: National Academy Press.
- Spybrook, J., Raudenbush, R.W., Liu, X., and Congdon, R. (2006). *Optimal Design for Longitudinal and Multilevel Research: Documentation for the "Optimal Design" Software*. Ann Arbor, MI: University of Michigan.
- SRA/McGraw-Hill (2003a). *DLM Early Childhood Express*. Desoto, TX: Author.
- SRA/McGraw-Hill (2003b). *Open Court Reading Pre-K*. Desoto, TX: Author.
- Stanovich, K.E. (1986). Matthew Effects in Reading: Some Consequences of Individual Differences in the Acquisition of Literacy. *Reading Research Quarterly*, 21: 360-407.
- Stone, J. (2002). *Animated Literacy*. La, Mesa, CA: Author. <http://www.animated-literacy.com>.
- Success for All Foundation. (2003a). *Curiosity Corner*. Baltimore: Author.
- Success for All Foundation. (2003b). *Kinder Corner*. Baltimore: Author.

References

- Swallow, W.H., and Monahan, J.F. (1985). Monte Carlo comparison of ANOVA, MIVQUE, REML, and ML estimators of variance components. *Technometrics*, 28: 47-57.
- U.S. Department of Education, National Center for Education Statistics. (2006). Enrollment in Early Childhood Education Programs. *The Condition of Education 2006* (NCES 2006-071). Washington, DC: U.S. Government Printing Office.
- U.S. Department of Health and Human Services. (2002). *A Descriptive Study of Head Start Families: FACES Technical Report I*. Washington, DC: Author.
- Verbeke, G., and Molenberghs, G. (2000). *Linear Mixed Models for Longitudinal Data*. New York: Springer.
- Wagner, R.K., Torgeson, J.K., and Rashotte, C.A. (1994). Development of Reading-related Phonological Processing Abilities: New Evidence of Bidirectional Causality from a Latent Variable Longitudinal Study. *Developmental Psychology*, 30: 73-87.
- Wagner, R.K., Torgeson, J.K., and Rashotte, C.A. (1999). *Comprehensive Test of Phonological Processing*. Austin, TX: Pro-Ed.
- West, J., Denton, K., and Germino-Hausken, E. (2000). *America's Kindergarteners* (ECLS-K) (NCES 2000-070). U.S. Department of Education. Washington, DC: National Center for Education Statistics.
- West, J., Denton, K., and Reaney, L. (2001). *Children's Reading and Mathematics Achievement in Kindergarten and First Grade* (NCES 2002-125). U.S. Department of Education. National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.
- Williams, K.T., and Wang, J.J. (1997). *Technical References to the Peabody Picture Vocabulary Test (PPVT-III)*. Circle Pines, MN: American Guidance Service, Inc.
- Wolfinger, R., Tobias, R., and Sall, J. (1994). Computing Gaussian Likelihoods and Their Derivatives for General Linear Mixed Models. *SIAM Journal on Scientific Computing* 15(6): 1294-1310.
- Wood, J. (2002). *Early Literacy and Learning Model*. Jacksonville, FL: Florida Institute of Education and the University of North Florida.
- Wright Group, McGraw Hill. (2001). *Doors to Discovery: A New Prekindergarten Program*. Bothell, WA: WrightGroup/McGraw-Hill.
- Zins, J., Bloodworth, M., Weissberg, R., and Walberg, H. (2004). The Scientific Base Linking Social and Emotional Learning to School Success. In J. Zins, R. Weissberg, M. Wang, and H. J. Walberg (Eds.), *Building Academic Success on Social and Emotional Learning: What Does the Research Say?* (pp. 1-22). New York: Teachers Press, Columbia University.