

**Longitudinal Study of Georgia's
Pre-K Program
Third Grade Report
January 2021**



**Georgia Dept
of Early Care
and Learning**
BRIGHT FROM THE START



FRANK PORTER GRAHAM
CHILD DEVELOPMENT INSTITUTE



We wish to acknowledge the members of our Georgia's Pre-K Program Evaluation Team who assisted with the 2017-2018 study, the previous study PI, Dr. Ellen Peisner-Feinberg; the Research Staff: Dr. Karen Van Manen, Andrea Reubens, Jada Walker, Dr. Treshawn Anderson, Ivonne Carrillo, Shannon Casey, Stephanie Catlett, Elizabeth Gunn, Cristina Luna Evans, Diana Knechtel, Alexis Kokoska, Tom Leggett, Carla Moreno, Jennifer Osborne, Judy Owen, Maria Sanchez, Eric Savage, Stefanie Schwartz, Gina Walker; and Data Collectors Brita Alton, Jane Argo, Joan Boudousquie, Elizabeth Bragg, Brianna Castellano, Cynthia Coleman, Ekeshia Crum, Kathy Crumley, Abby Davis, Akua Ducard, CL Dunn, Deborah Hallman, Stacey Harris, Khadija Highsmith, Melissa Johns, Deborah Laing, Kim Lane, Leslie Lane, LaJuana Johnson-Mendez, Elizabeth Milling, Ann O'Mahoney, Chigozie Nkemka, Jacqueline Parker, Nicole Poucher, Katie Rutland, Carolina Sandoval, Steven Schell, Melissa Silva, Patricia Simpson, Moneesha Smith, Karen Stewart, Alissa Sypsa, Anna Tussey, Peggy Tymes, Lori Warren, Adriana Vivas-Sosa, Ann Zimmer-Shepherd, Heather Zook; and the previous analysis team: Irina Mokrova and Margaret Burchinal.

In addition, we share our appreciation to all those who participated in and assisted with this study, including the children and families; the teachers, administrators, and other staff of Georgia's Pre-K Program and the kindergarten, first, second, and third grade classrooms and schools; and the staff of Bright from the Start: Georgia Department of Early Care and Learning (DECAL).

Cover by Jennifer Osborne.

This study was funded by Bright from the Start: Georgia Department of Early Care and Learning. The opinions expressed in this report do not necessarily reflect those of the funding agency.

Suggested citation: Soliday Hong, S., Zadrozny, S., Walker, J., Love, E.N.G., Osborne, J.D., Owen, J. L., Jenkins, G., & Peisner-Feinberg, E. (2021). *Longitudinal Study of Georgia's Pre-K Program: Third Grade Report*. Chapel Hill, NC: The University of North Carolina, FPG Child Development Institute.

This report is available at <http://fpg.unc.edu/projects/georgia-pre-kindergarten-evaluation> or at <http://www.decal.ga.gov/BftS/EvaluationGAPreKProgram.aspx>.

Table of Contents

List of Tables	4
List of Figures.....	4
Executive Summary	5
Longitudinal Study of Georgia’s Pre-K Program	6
Overview of Georgia’s Pre-K Program	6
Previous Evaluation Studies of Georgia’s Pre-K Program	7
Approach	8
Recruitment	8
Participants.....	9
Child Assessments	11
Classroom Observations.....	13
Parent and Teacher Surveys.....	13
Third Grade Results	14
Comparison of Within Grade Scores with National Averages	15
Patterns of Growth in Children’s Scores Over Time	17
Patterns of Growth in Children’s Scores Over Time: DLL Subsample.....	19
Predictors of Patterns of Growth in Children’s Scores Over Time.....	21
Pre-K Program/Classroom Characteristics	25
Third Grade Longitudinal Study Summary	29
Comparison Sub-study	31
Comparison of Third Grade Scores: Georgia’s Pre-K versus No Pre-K	31
Comparison Group Recruitment.....	31
Parent and Teacher Surveys.....	31
Comparison Group Participants	32
Child Assessments	33
Comparison Group Analyses	33
Summary and Conclusions on the Comparison Sub-study	34
Appendix A. Child Outcome and Classroom Quality Measures	35
References.....	36

List of Tables

[Table 1.](#) Number of Classrooms and Children Participating in the Longitudinal Study

[Table 2.](#) Constructs/Measures and Years Administered in the Longitudinal Study

[Table 3.](#) Fall-Spring Gains by Grade – Effect Sizes

List of Figures

[Figure 1.](#) Longitudinal Study Timeline

[Figure 2.](#) Overview of Study Activities in Third Grade

[Figure 3.](#) Statewide Distribution of Participating School Districts in Pre-K and Third Grade

[Figure 4.](#) Children and Family Demographic Characteristics for the Third Grade Sample Compared to Child Who Participated in Pre-K but Not Third Grade (Attrition)

[Figure 5.](#) Fall-Spring Gains by Grade – Rate of Growth within School Years

[Figure 6.](#) Growth in Language and Literacy Skills (Pre-K – 3rd Grade)

[Figure 7.](#) Growth in Math Skills (Pre-K – 3rd Grade)

[Figure 8.](#) Growth in Executive Function Skills (1st – 3rd Grade)

[Figure 9.](#) Growth in Social Skills (Pre-K – 3rd Grade)

[Figure 10.](#) Pre-K and First Grade versus Third Grade Scores

[Figure 11.](#) Rate of DLL Growth in English Language and Literacy Skills

[Figure 12.](#) Rate of DLL Growth in Spanish Language and Literacy Skills

[Figure 13.](#) DLL Growth in Math Skills Assessed in English

[Figure 14.](#) DLL Growth in Math Skills Assessed in Spanish

[Figure 15.](#) DLL Growth in Executive Function Skills

[Figure 16.](#) Growth in Picture Vocabulary Skills Over Time by Pre-K English Language Proficiency Level

[Figure 17.](#) Growth in Applied Problem Skills Over Time by Pre-K English Language Proficiency Level

[Figure 18.](#) Growth in Social Skills Over Time by Pre-K English Language Proficiency Level

[Figure 19.](#) Growth in Vocabulary Skills Over Time by IEP Status at the Beginning of Pre-K

[Figure 20.](#) Growth in Problem Behaviors Over Time by Sex of Child

[Figure 21.](#) Growth in Letter-Word Identification Skills Over Time by Kindergarten-Third Grade Classroom Quality: Low/Moderate versus High

[Figure 22.](#) Growth in Word Attack Skills Over Time by Kindergarten-Third Grade Classroom Quality: Low/Moderate versus High

[Figure 23.](#) Growth in Problem Behaviors Over Time by Kindergarten-Third Grade Classroom Quality: Low/Moderate versus High

[Figure 24.](#) Average CLASS Total Scores by Grade

[Figure 25.](#) Frequency of CLASS Average Total Scores by Grade

[Figure 26.](#) Frequency of CLASS Average Total Scores by Domain: Emotional Support

[Figure 27.](#) Frequency of CLASS Average Total Scores by Domain: Classroom Organization

[Figure 28.](#) Frequency of CLASS Average Total Scores by Domain: Instructional Support

[Figure 29.](#) Third Grade Comparison Group versus Original Pre-K Group

[Figure 30.](#) Differences in Standard and Raw Scores for Pre-K Group versus Comparison Group in Third Grade

Executive Summary

The following summarizes results from the Georgia's Pre-K Longitudinal Study Third Grade Report. The study followed a representative sample of children (n=1,169) from their pre-k experience (2013-2014) through the third grade (2017-2018). The study was conducted by researchers at the Frank Porter Graham (FPG) Child Development Institute at the University of North Carolina at Chapel Hill.

Study Components:

- Standardized child assessments that measure skills across learning domains. By using standardized child assessments, children's scores can be compared to a nationally representative sample of children of the same age;
- Observations of classroom quality over time that measure the quality of teacher-child interactions;
- Inclusion of a subsample of dual language learners (English-Spanish); and
- A comparison sample of children who did not attend any pre-k added in the third grade year.

Key Results of the Longitudinal Study:

- The largest gains (higher than expected scores relative to the norming sample) were observed during the Georgia's Pre-K year. Small to moderate gains were found in measures assessing literacy, math, and social skills;
- Initial gains persisted through Kindergarten with scores starting to level off in first grade and then decreasing or stabilizing through third grade.
- For four of the nine assessments, children's scores were higher in third grade than at the pre-k baseline;
- At the end of third grade, children's scores, on average, were below the national norm on two of the measures: Vocabulary and Passage Comprehension. Vocabulary scores were below the national norm throughout the study;
- Scores for children classified as Dual Language Learners (DLL) were slightly below the national norm in third grade for skills measured in English and well below the national norm for the skills measured in Spanish;
- The strongest predictor of children's scores over time is English Language Proficiency at pre-k entry. Predictors associated with small differences in a few outcomes included: Individualized Education Plan (IEP) status at the beginning of pre-k, children's sex, and K-3 classroom quality; and
- Classroom quality, as measured by the Classroom Assessment Scoring System (CLASS) Pre-K and K-3 versions, was in the moderate range, on average, and highest in Pre-K but only declined slightly in later grades.

Key Results from the Comparison Sample:

Children who attended Georgia's Pre-K Program had literacy skills that were moderately higher and executive function skills that were somewhat higher in the fall of third grade than children whose parents reported that the child did not attend any pre-k program (comparison group). These results are similar to the findings of the Longitudinal Study where children who attended Georgia's Pre-K had higher scores in foundational literacy skills relative to the national norming sample. Together, these results suggest that foundational literacy skills, which are a focus of pre-k, were not obtained by children in the comparison sample.

Longitudinal Study of Georgia's Pre-K Program Third Grade Report

The purpose of this study was to examine associations between attendance in Georgia's Pre-K Program and children's academic and social outcomes through third grade. Different from previous years' reports, this includes a sub-study that incorporates a comparison group of children who did not attend any pre-k program. This report focuses on third grade outcomes, but it also includes pre-k to third grade outcomes collected during the 2013–2014 to 2017–2018 school years and a comparison sub-study of children who did not attend any pre-k.

The longitudinal study began with a sample of 1,169 children (139 Spanish speaking dual language learners/DLLs) attending a random sample of 199 Georgia's Pre-K classes in the first year of the study. Eight-hundred fifty-seven children (109 Spanish-speaking DLLs) were followed into third grade. Professional assessors conducted individual child assessments near the beginning and end of each school year. The assessment measured multiple areas of learning, including language, literacy, math, executive function, and teacher ratings of behavior skills. For the DLL subsample, parallel assessments were conducted in English and Spanish. Researchers also conducted observations in all pre-k classrooms and a subset of children's classrooms each year from kindergarten through third grade. In addition, parents and teachers were surveyed, and some administrative pre-k data were used.

Overview of Georgia's Pre-K Program

Georgia's Pre-K Program is a state-funded, universal prekindergarten program for four-year olds from all income levels that serves more than 80,000 children each year. Georgia's Pre-K Program was established in 1992, and Georgia became one of the first states to offer a universal program in 1995. The program is administered by Bright from the Start: Georgia Department of Early Care and Learning (DECAL). The following are key components of the program:

- Variety of settings: public school systems, private providers, and blended Head Start/pre-k classrooms
- Based on a school-year model with instruction for 180 days/year and 6.5 hours/day^a
- Class sizes limited to 20–22 children with a lead and assistant teacher, and adult: child ratios of 1:11
- Lead teachers required to have at least a bachelor's degree in early childhood education or a related field or a bachelor's degree in any field along with an approved early childhood education credential
- Assistant teachers required to have at least a Paraprofessional Certificate (issued by the Georgia Professional Standards Commission) or a Child Development Associate (CDA) credential

^a Before 2011–2012, Georgia's Pre-K Program provided 180 instruction days per year, but budget restrictions led to a reduction to 160 days in 2011–2012. In 2012–2013, the program year was increased to 170 days, and in 2013–2014, it was returned to 180 days.

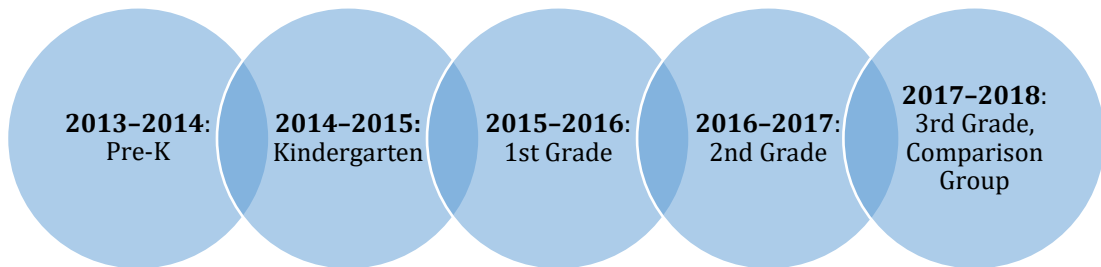
- Funding provided by the Georgia Lottery for Education provides minimum salary requirements for lead teachers and assistant teachers based on credentials

Previous Evaluation Studies of Georgia’s Pre-K Program

In 2011, the Georgia legislature funded a series of studies to evaluate the impact of Georgia’s Pre-K Program on children’s outcomes over time.

1. **Pre-K Outcomes Study**, conducted in 2011–2012, examined children’s learning outcomes during pre-k, the factors that predicted better short-term outcomes, and the quality of children’s experiences in Georgia’s Pre-K classes. This study included 509 children recruited from a random sample of 100 pre-k classrooms. On average, children, including DLLs, exhibited significant growth during their pre-k year across all domains of learning, including language and literacy skills, math skills, general knowledge, and behavioral skills. Classroom quality was in the moderate to high range.
2. **Kindergarten Comparison Study**, conducted in 2012–2013, was a quasi-experimental design (QED) study that investigated the effects of participating in Georgia’s Pre-K Program on children’s school readiness skills compared to eligible children who had not yet attended the program. This QED study used a regression discontinuity design (RDD) and included 1,181 children (611 children who attended pre-k, and 570 children who had not yet attended pre-k). Children who attended Georgia’s Pre-K Program had better language and literacy, math, and general knowledge skills during kindergarten than children who did not attend. No differences were observed between groups on social skills or behavior problems.
3. **Longitudinal Study**, occurring from 2013–2021 has followed a sample of 1,169 children from pre-k through third grade, to examine the short- and long-term learning outcomes for children who attended Georgia’s Pre-K Program and to determine the quality of their preschool and school experiences (see [Figure 1](#) below). Previous reports from this study showed that children had gains in language, literacy, math, general knowledge, and social skills during their pre-k, kindergarten, and first grade years. For a few measures (math problem solving, vocabulary, reading comprehension), scores decreased slightly during the first-grade year and again in the second-grade year. For more advanced reading and math skills assessed only in first and second grades, standardized scores also showed decreases over time with children continuing to score in the expected range for their age. Results through third grade are reported below.

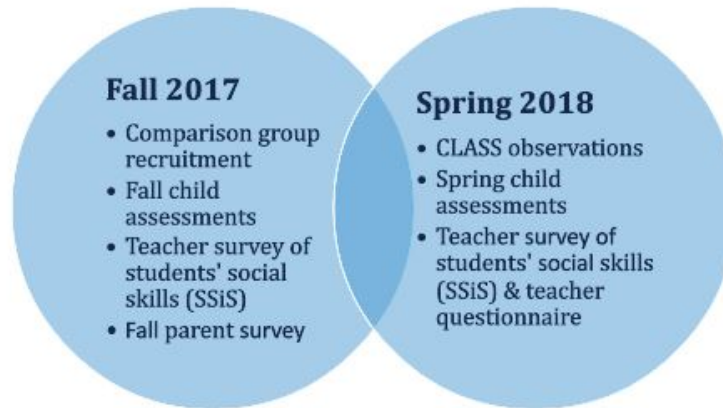
Figure 1. Longitudinal Study Timeline



Approach

This third grade report includes data collected with a longitudinal cohort of children from Georgia’s Pre-K Program through third grade. [Figure 2](#) provides an overview of the study activities in the children’s third grade year. The following sections describe the technical details related to participant recruitment and data collection procedures.

Figure 2. Overview of Study Activities in Third Grade



Recruitment

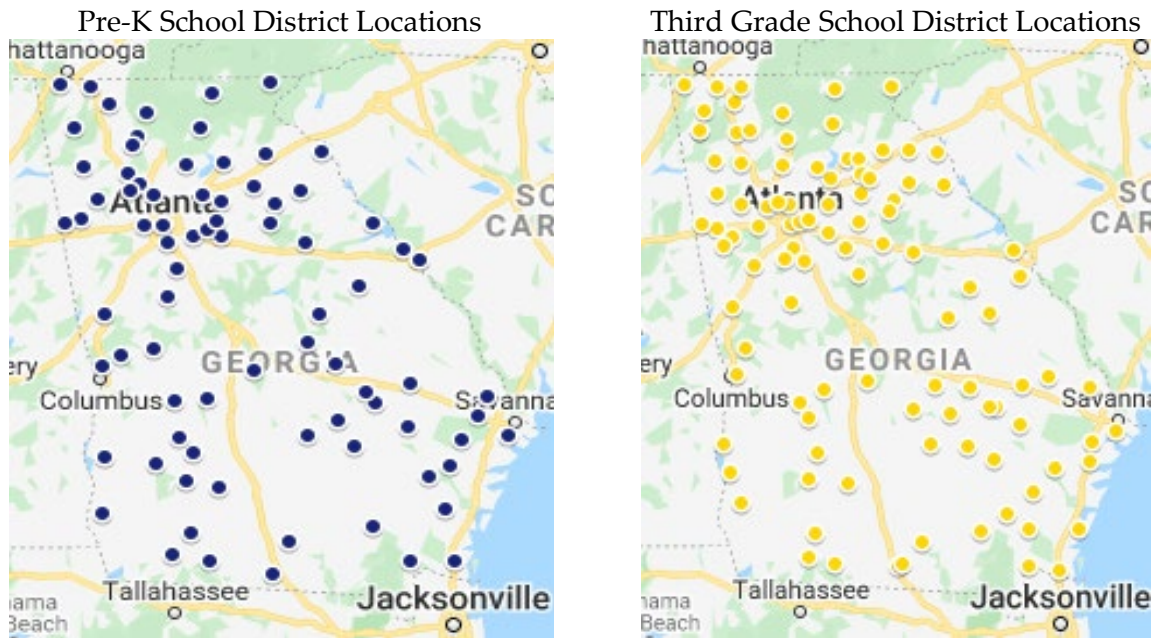
Children

The original longitudinal sample of children who attended Georgia’s Pre-K Program were recruited in the first year of the study (2013–2014). Parent permission forms were distributed to all children in 199 randomly selected Georgia’s Pre-K classrooms, with an overall permission rate of 73% (3,136 of 4,270 eligible children). From all returned permission forms, an average of six children per classroom were randomly selected for inclusion in the study.

Districts, Schools, and Teachers

Every fall, the research team recruited school districts, principals, and teachers where study participants attended school after pre-k. If the school district and the school principal did not grant approval for researchers, the participating student was not assessed. In third grade, the research team obtained approval from 105 of 114 district superintendents for school districts where study participants attended third grade. From those districts, 432 school principals granted approval for schools and teachers to participate in the third grade study. Of the 654 classrooms in which study children were enrolled, 640 teachers gave consent, and 14 refused participation. Participating schools were located throughout Georgia including rural and urban areas. See [Figure 3](#) for a map of the districts that participated in pre-k and third grade (charter and private schools are not included to protect anonymity of the study participants).

Figure 3. Statewide Distribution of Participating School Districts in Pre-K and Third Grade



Participants

Teachers and Classrooms

Children in the initial year of the longitudinal study (2013–2014) attended 199 Georgia’s Pre-K classrooms; in 2014–2015, 822 kindergarten classrooms; in 2015–2016, 777 first-grade classrooms; in 2016–2017, 786 second-grade classrooms; and in 2017–2018, 718 third grade classrooms. See [Table 1](#). About half of the pre-k classrooms attended by children in the study sample were in public school settings (49%), and about half were in private settings (51%). As children were followed over time, the elementary school classrooms attended by children in the study were primarily located in public school settings (3rd= 98%), with a few in charter schools and private schools (3rd=1%).

Table 1. Number of Classrooms and Children Participating in the Longitudinal Study

Study Year	Total Classrooms	CLASS Observations	Total Children	DLLs
Pre-K (2013–2014)	199	199	1,169	139
Kindergarten (2014–2015)	822	196	1034 (88%)	118 (85%)
Grade 1 (2015–2016)	777	296	969 (83%)	119 (86%)
Grade 2 (2016–2017)	786	280	951 (81%)	116 (83%)
Grade 3 (2017–2018)	718	256	857 (73%)	109 (78%)

Note: Percentages noted in parentheses are the proportion of the pre-k sample assessed in each grade.

Classroom and teacher characteristics were collected each year through teacher surveys and state administrative data from DECAL in pre-k. The average class size was similar each year of the study, about 20–21 children, with half boys and half girls. The majority (82%) of pre-k teachers were certified by the Georgia Professional Standards Commission. In pre-k, almost two-thirds (64%) of the teachers had a bachelor's degree and about one-third (34%) had a master's degree or higher. In third grade this pattern was reversed, with slightly fewer than one-third having a bachelor's degree (32%) and more than two-thirds having a master's degree or higher (68%). Third grade teachers reported a substantial number of years of teaching experience (13 years on average) with more than six years teaching at their current grade level. In pre-k, teachers reported 11 years teaching and six in their current grade level. The majority of third grade teachers were White (73%; PK 67%); almost one-fifth were Black (19%; PK 27%); a small proportion were of Hispanic/Latino ethnicity (3%; PK 3%); nearly all were female (95%; PK 98%).

Children

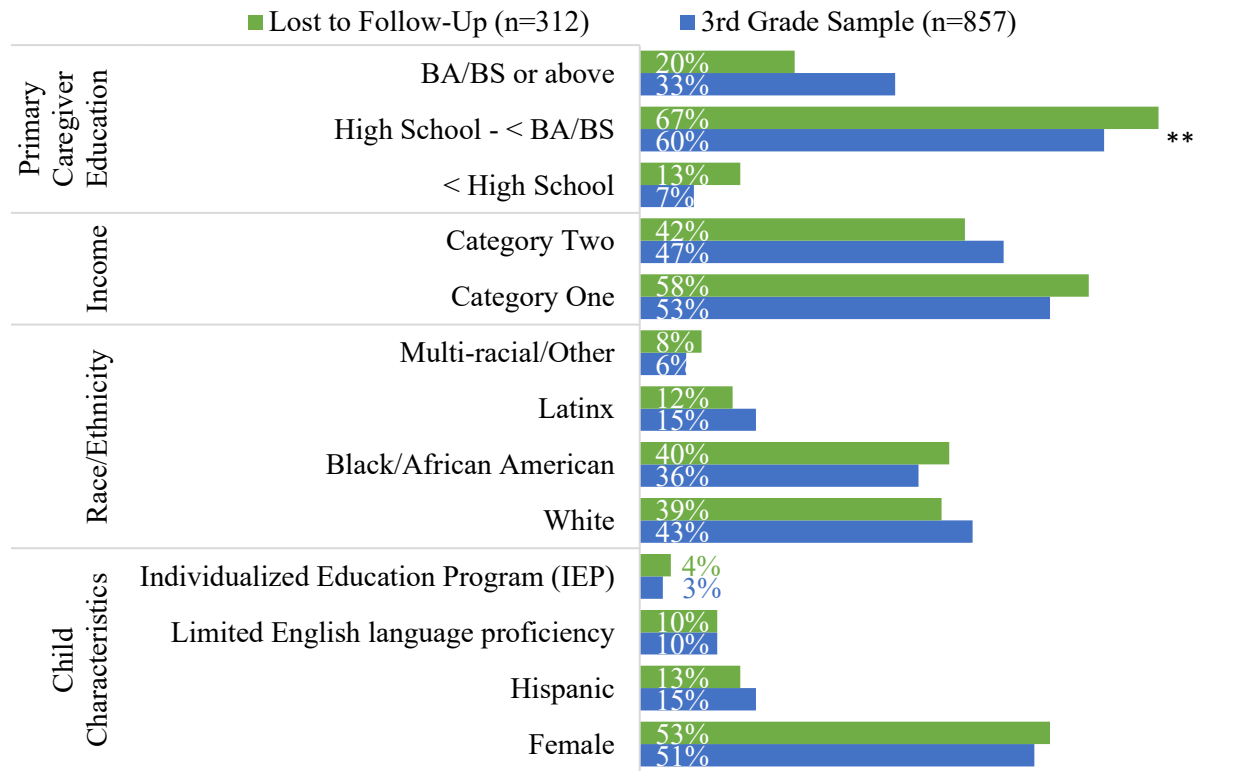
The retention rate in the third grade sample was strong. The original pre-k sample included 1,169 children; the third grade sample included 857 children (73% of the original sample). These children included a subsample of Spanish-speaking DLLs: 139 children in the original sample and 109 children in third grade (78% of the original sample). Previous studies of other pre-k programs have had similar or lower retention rates (e.g., New Jersey Abbott Pre-K follow-up had an overall retention rate of 70% in third grade ⁱ; 65% overall retention rate in Tulsa Pre-K study at third grade ⁱⁱ; and 36% retention rate in the Tennessee Pre-K Study at third gradeⁱⁱⁱ).

Information about child and family characteristics for the study sample in third grade was obtained from parent survey data (see [Figure 4](#)). About half the children in the third grade sample were boys (49%); about half were girls (51%). The children were from varied racial/ethnic backgrounds, including 43% White, 36% Black/African American, 15% Latino, and the remainder from other or multiracial backgrounds (6%). Approximately 15% of children were Hispanic. Ten percent of the children had limited English language proficiency, and 3% had an individualized education plan (IEP) at the beginning of pre-k. Slightly more than half (53%) of the children were from low-income families as indicated by Pre-K Category One status (which represented participation in one or more programs including SNAP, TANF, SSI, CAPS, Medicaid, and free or reduced-price meals). Students with Category Two status were not income-eligible for the programs listed above. The education level for the majority (60%) of children's primary caregivers was between a high school diploma and less than a bachelor's degree, with 33 percent having a bachelor's degree or above.

An attrition analysis was conducted to compare child and family characteristics of the children who were assessed in third grade and the original pre-k sample to examine whether the two groups were equivalent on these characteristics (see [Figure 4](#)). Children in the third grade

sample had more parents with higher rates of bachelor’s degree or above.^a In every other area, the two groups were demographically similar. In order to ensure that the results of subsequent analyses generalize to the full sample, analyses in this report include characteristics where the third grade sample differed from the original sample as covariates (primary caregiver education level).

Figure 4. Child and Family Demographic Characteristics for the Third Grade Sample Compared to Children Who Participated in Pre-K, but Not Third Grade (Attrition)



* p-value < .05, ** p < .10, *** p < .001

Note: Category One status represented participation in one or more programs including SNAP, TANF, SSI, CAPS, Medicaid, and free or reduced-price meals. Students with Category Two status were not income-eligible for the programs listed above.

Child Assessments

Individual assessments used to measure children’s skill growth were conducted in children’s classrooms each year. Children who attended Georgia’s Pre-K Program were assessed in English at 10 time points, in the fall and spring near the beginning and end of each year from pre-k through third grade. In addition to parental permission, children’s verbal assent was obtained before the assessments began. Children whose home language was Spanish also received a set of parallel assessments using Spanish language versions of these measures. The

^a Results reflect standard comparisons with Absolute Standardized Differences (Hodges-Lehmann Estimator) between the full longitudinal pre-k sample and the children assessed in third grade.

assessments in Spanish were conducted by a different, bilingual data collector approximately two weeks after the assessments in English.

Measures of language, literacy, and math skills were gathered from pre-k through third grade using the Woodcock-Johnson III: Tests of Achievement (Woodcock, McGrew, & Mather, 2001) and the Bateria III Woodcock-Muñoz: Pruebas de aprovechamiento (Bateria III; Muñoz-Sandoval, Woodcock, McGrew, & Mather, 2005). The forward and backward digit span tasks were adapted from Gathercole & Pickering (2001). In first grade, more developmentally advanced measures of reading comprehension, number operations, and executive function were added to the battery (see [Table 2](#) and [Appendix A](#) for a list of measures and years of administration). All the child assessment measures were available in English and Spanish versions. Most of the measures used were norm-referenced, and therefore, standard scores could be used for these measures. These scores consider children’s ages, so the standardized mean score of 100 represents the expected performance for an average child at a given age.

Table 2. Constructs/Measures and Years Administered in the Longitudinal Study

Pre-K	K	1st	2nd	3rd
Language: PreLAS 2000	→			
		Executive Function: Digit Span (Forward & Backward) in English and Spanish*	→	
Reading-Decoding: WJ/Bateria - Letter Word ID	→			
Language: WJ/Bateria - Picture Vocabulary	→			
Reading-Phonetic Decoding: WJ/Bateria - Word Attack	→			
Reading-Phonetic Decoding: WJ/Bateria - Sound Awareness	→			
		Reading Comprehension: WJ/Bateria - Passage Comprehension*	→	
Math Reading: WJ/Bateria - Applied Problems	→			
		Math Computation: WJ/Bateria - Calculation*	→	

Note: * Indicates measures that were first administered in first grade

Classroom Observations

Observations of classroom practices were conducted in classrooms attended by children in the sample each year. When the children were in pre-k, observations were conducted in all 199 randomly selected pre-k classrooms. In kindergarten, a random sample of classrooms was selected for observation. In subsequent years, the children enrolled in classrooms that were observed in kindergarten were followed, and their classrooms were observed to obtain a longitudinal sample associated with the classroom observations. In third grade, observations were conducted in 256 of the 706 third grade classrooms attended by children in the sample (representing 496 children in the longitudinal sample). Classroom observations were conducted during the second half of the school year each year. Observations typically lasted two to three hours. Data collectors completed standard training procedures offered by the developers of the measures, had additional field practice, and had to meet established reliability criteria before gathering data.

The classroom observations used the same measure each year to examine the quality of teacher-child instructional interactions with appropriate versions for the age range of children. The *Classroom Assessment Scoring System (CLASS) Pre-K* version (Pianta, La Paro, & Hamre, 2008) was used in pre-k and the *CLASS K-3* version (Pianta, La Paro, & Hamre, 2008) was used in kindergarten through third grade. The *CLASS Pre-K/CLASS K-3* measured teachers' interactions with children regarding social and emotional interactions, classroom organization and management, and instructional interactions with children. Both versions of the *CLASS* include three domains: Emotional Support, Classroom Organization, and Instructional Support. Total scores were also calculated.

Parent and Teacher Surveys

Parents completed demographic surveys each fall (pre-k to third grade) about their family and household. The annual fall parent survey included information about parent education levels, income, languages spoken, and number of adults/children in the home. Parent surveys were received from 91% (1,067/1,169) of participating families in pre-k. Surveys received in third grade represented 71% (609/857) of participating families.

Teachers completed surveys in the spring that included information about their classrooms and their background, including classroom composition (e.g., number of boys and girls in class and languages spoken by the children); years of teaching experience and degrees earned; available classroom resources; level of perceived control over what is taught and how; and levels of parent involvement and support. Teacher survey completion rates were high--95% (189/199) in pre-k and 84% (590/699) in third grade.

Third Grade Results

Third grade results are organized into five areas: (1) comparison of within grade scores with national averages, (2) patterns of growth in children's scores over time from pre-k through third grade in English and Spanish, (3) predictors of growth in children's scores over time, (4) examination of classroom quality, and (5) comparison of third grade scores for children who attended Georgia's Pre-K Program with children who did not attend any pre-k program.

Analytic Approach: Longitudinal growth models were estimated from pre-k through third grade for children who attended Georgia's Pre-K Program using two-level hierarchical linear models (HLM), with time at level one and child at level two. This approach allowed for examination of patterns of skill growth from pre-k to third grade, accounting for factors known to be related to school performance such as child, family, program, and classroom characteristics and quality. Growth over time was allowed to follow a non-linear pattern, and the HLM growth models included linear, quadratic, and cubic terms where appropriate. All outcomes were modeled as cubic functions except the SSiS Problem Behavior Rating Scale, which followed a quadratic functional form, and the Backward Digit Span task, which was linear. The models included two sets of covariates:

- 1) Child/family characteristics – sex of child, race/ethnicity (Hispanic/Latino, White, Black, multi-racial/other), IEP status at the beginning of pre-k, English/Spanish preLAS language proficiency level in fall of pre-k, and family income category.
- 2) Pre-K program/classroom characteristics – provider type (private setting versus public school system) and class size; pre-k classroom quality, and subsequent kindergarten through third grade classroom quality.

Within grade scores were estimated from the regression models above as marginal means.

Standard Scores: Standard scores were used to compare children's expected scores accounting for typical growth for age over time. Without intervention, it is expected that children's scores would remain fairly stable over time, a change of zero. A change greater than zero, or growth, indicates that children gained additional skills in that skill relative to the amount of growth of other children of the same age in that skill area. Negative change over time indicates that children did not make as much growth in the skill area as same-age peers, often having average scores that trend downward toward the average score for children of the same age. Standard scores are available for all assessments except executive function measured by the Forward and Backward Digit Span tasks. Standard scores have a national norm mean of 100 and a standard deviation of 15.

Comparison of Within Grade Scores with National Averages

Within grade gains in standard effect sizes are shown in [Table 3](#) for measures assessed using standard scores. [Figure 5](#) shows the same differences in standard scores with the addition of executive function scores in their original scale. The largest gains were observed during the pre-k year where, from fall to spring, children made small to moderate gains in two domains of literacy and social skills. Small gains were also observed in kindergarten related to auditory processing and decoding skills along with very small gains in letter-word identification. During second grade, children’s literacy and math skill growth was smaller relative to the national average (negative growth), but the rate of decline was in the very small range and resulted in children’s scores moving closer to the national average. Finally, in third grade, children’s growth in their auditory processing and letter-word identification skills slowed from fall to spring but showed a small amount of growth in their social skills relative to peers of the same age.

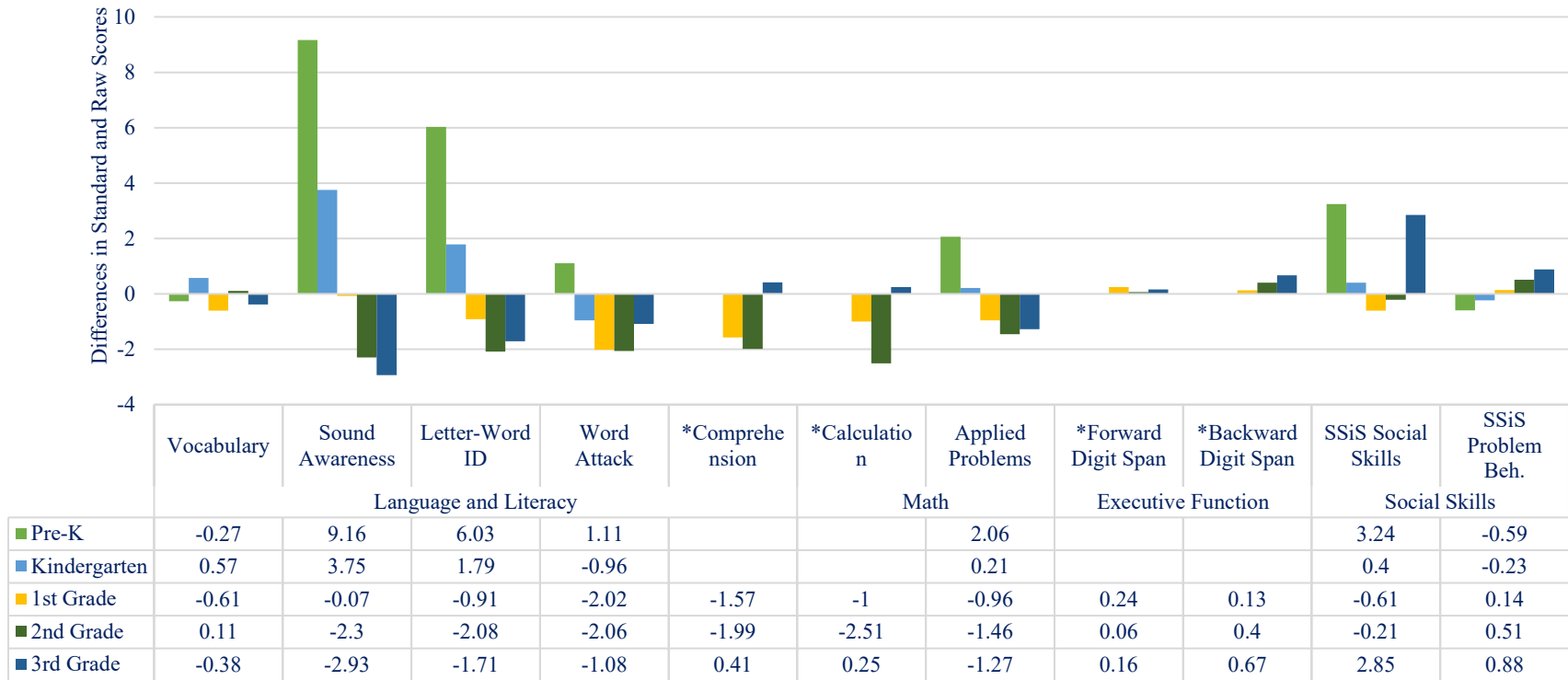
Table 3. Fall-Spring Gains by Grade – Effect Sizes

		Pre-K	Kindergarten	1st Grade	2nd Grade	3rd Grade
Language & Literacy	Picture Vocabulary	-0.02	0.04	-0.04	-0.03	0.01
	Sound Awareness	0.61	0.25	0.00	-0.15	-0.20
	Letter-Word ID	0.40	0.12	-0.06	-0.14	-0.11
	Word Attack	0.07	-0.06	-0.13	-0.14	-0.07
	Comprehension	X	X	-0.10	-0.13	0.03
Math	Calculations	X	X	-0.07	-0.17	0.02
	Applied Problems	0.14	0.01	-0.06	-0.10	-0.08
Social Skills	Social Skills	0.22	0.03	-0.04	-0.01	0.19

Note:

X Not administered that study year	Very small gains or losses from fall to spring within each year (E.S. = 0.10-0.19).	Small gains or losses from fall to spring within each year (E.S. = 0.20-0.49).	Moderate gains or losses from fall to spring within each year (E.S. = 0.50-0.79).
--	---	--	---

Figure 5. Fall-Spring Gains by Grade – Rate of Growth within School Years.



*=Not administered in pre-k and kindergarten.

Patterns of Growth in Children's Scores Over Time

For children who attended Georgia's Pre-K Program, average scores were within the expected and average range (standard score of 90-110), with some skills in the "high average" range (standard score of 111-120). Patterns of longitudinal growth were significant over time for most language, literacy, math, and social skills measured from pre-k through third grade assessed in English (see Figures 6, 7, 8, and 9). On most of these norm-referenced measures, children showed a pattern of initial gains in scores during pre-k and kindergarten (i.e., larger gains than expected relative to the norming sample), with scores starting to level off in first grade and then growth decreasing, or scores stabilizing, through third grade.

Figure 6. Growth in Language and Literacy Skills (Pre-K – 3rd Grade)

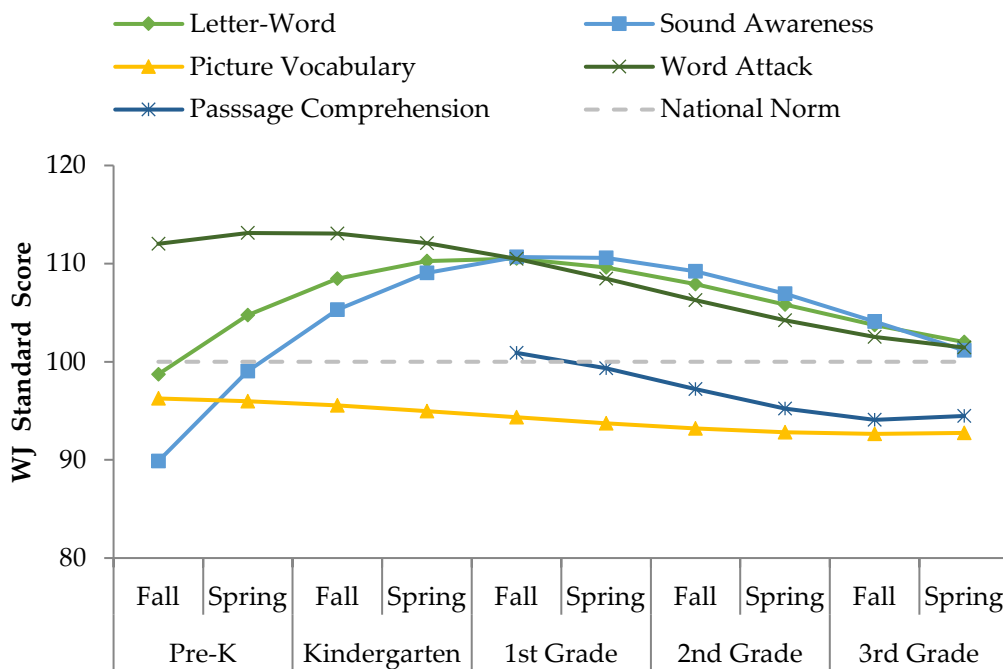


Figure 7. Growth in Math Skills (Pre-K – 3rd Grade)

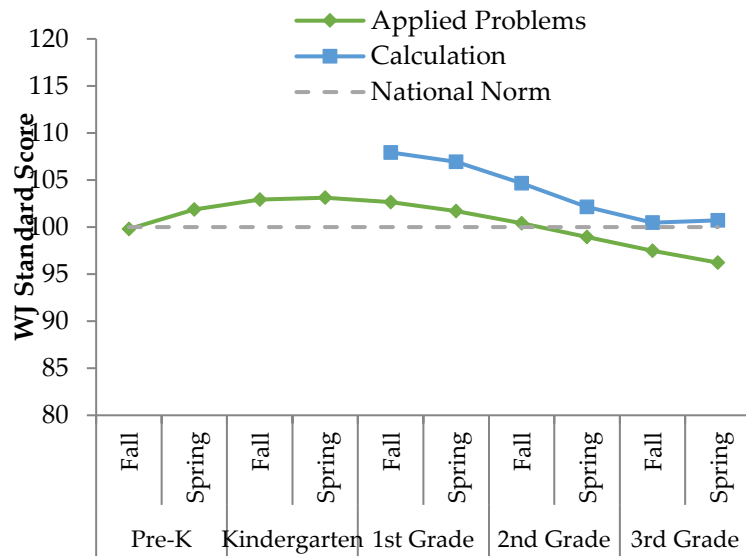
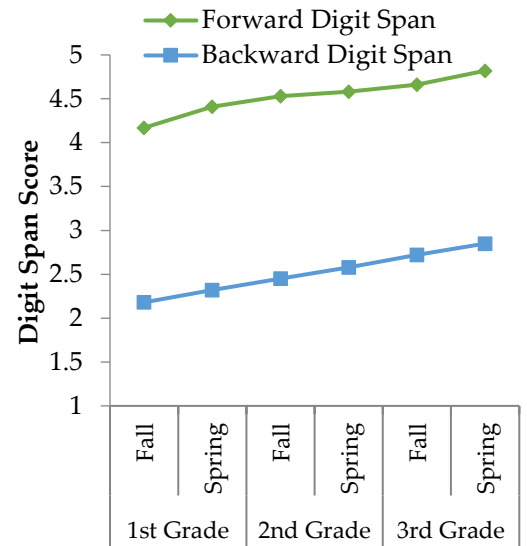
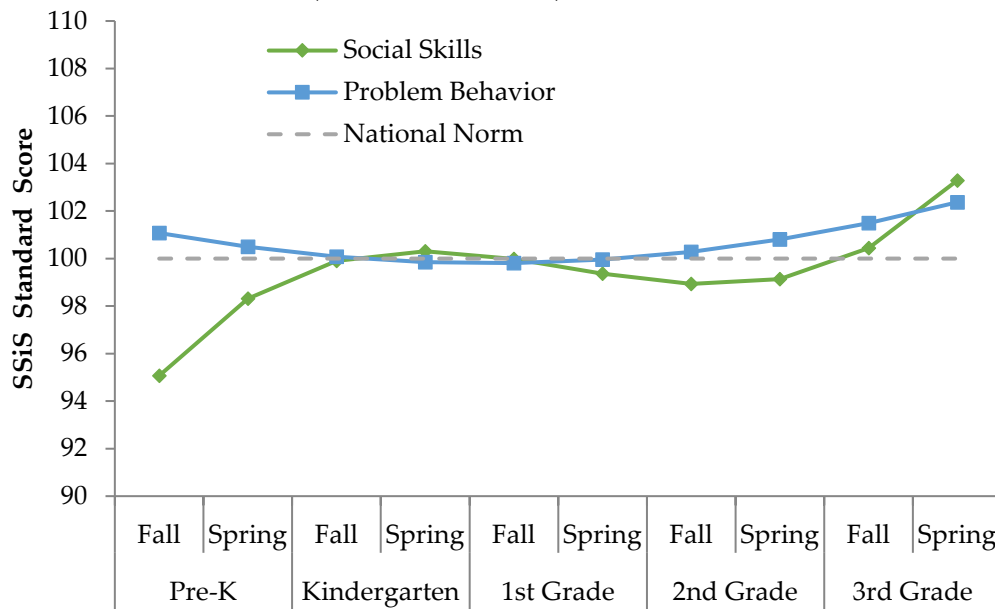


Figure 8. Growth in Executive Function Skills (1st – 3rd Grade)*



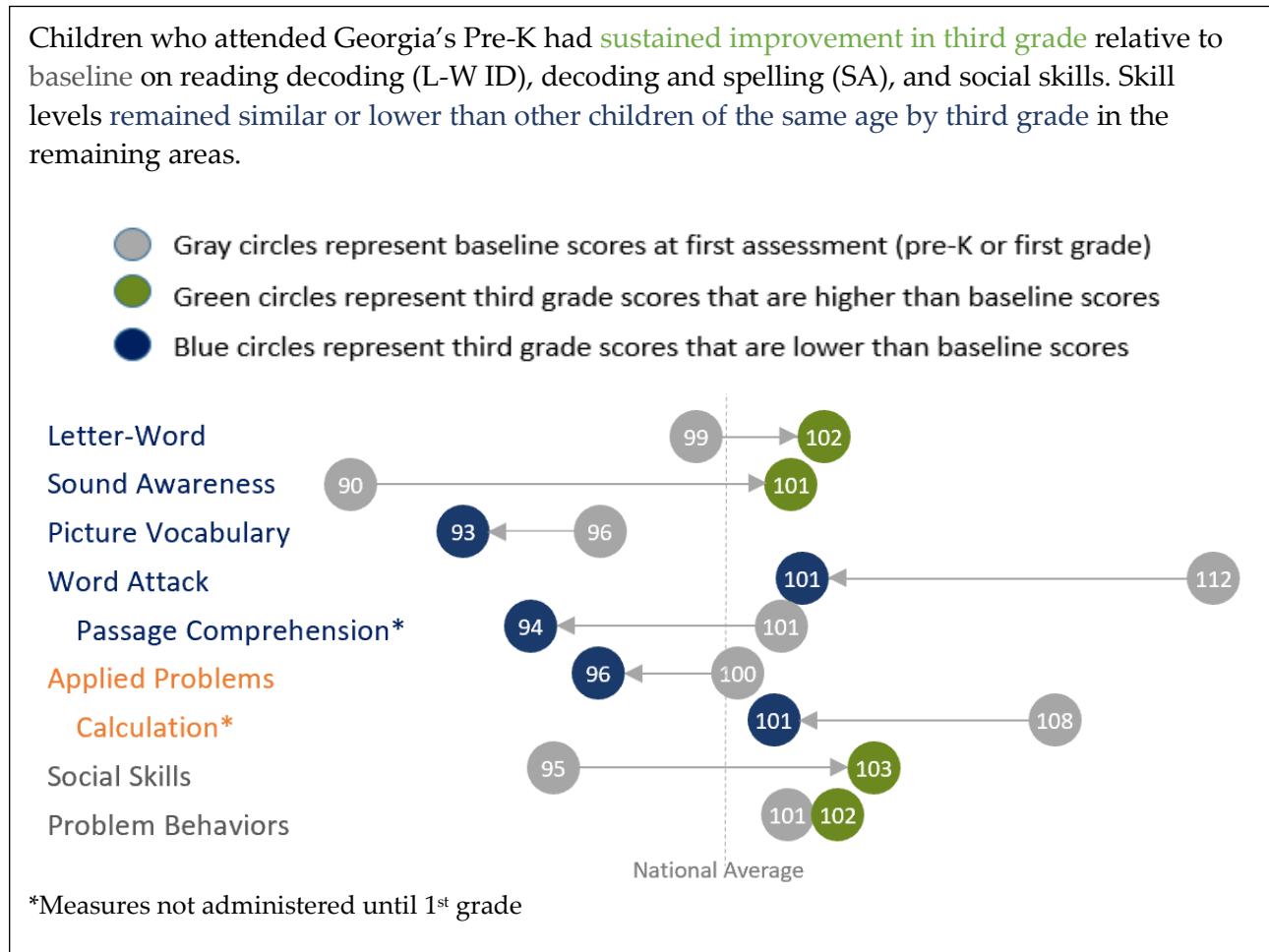
*Figure 8. Digit Span measures produce raw scores, not standardized scores - national averages are not available.

Figure 9. Growth in Social Skills (Pre-K – 3rd Grade)



Comparisons of baseline pre-k scores (or first grade for assessments that were first administered after pre-k) are found in [Figure 10](#).

Figure 10. Pre-K and First Grade versus Third Grade Scores



Patterns of Growth in Children’s Scores Over Time: DLL Subsample

Researchers conducted a parallel set of analyses specifically for dual language learners (DLLs). Like the full sample analyses, patterns of growth were examined over time in child skills from pre-k through third grade for children who attended Georgia’s Pre-K Program. Like the results for the full sample, children showed larger gains during pre-k and somewhat lesser gains during kindergarten, with smaller-than-expected gains during first grade and leveling out or greater decreases during second and third grade. However, compared to the full sample, the average scores were somewhat lower in English (see Figures [11](#), [13](#), and [15](#)) and much lower in Spanish for DLLs (see Figures [12](#) and [14](#)).

Figure 11. Rate of DLL Growth in English Language and Literacy Skills

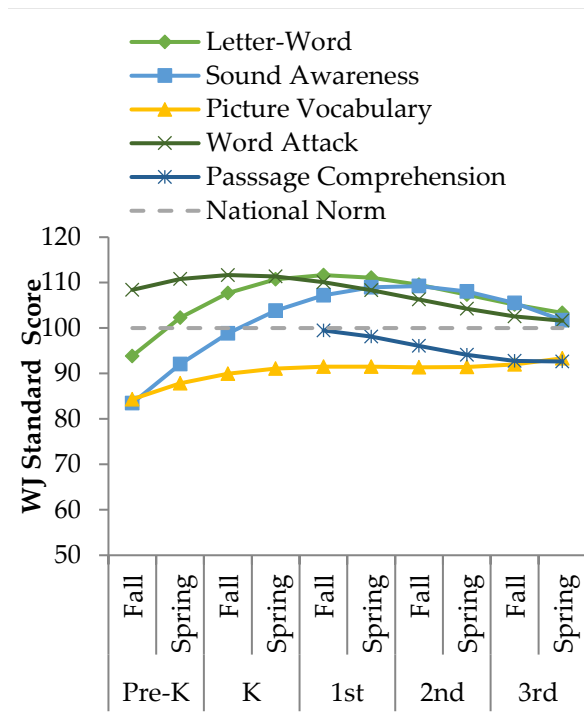


Figure 12. Rate of DLL Growth in Spanish Language and Literacy Skills

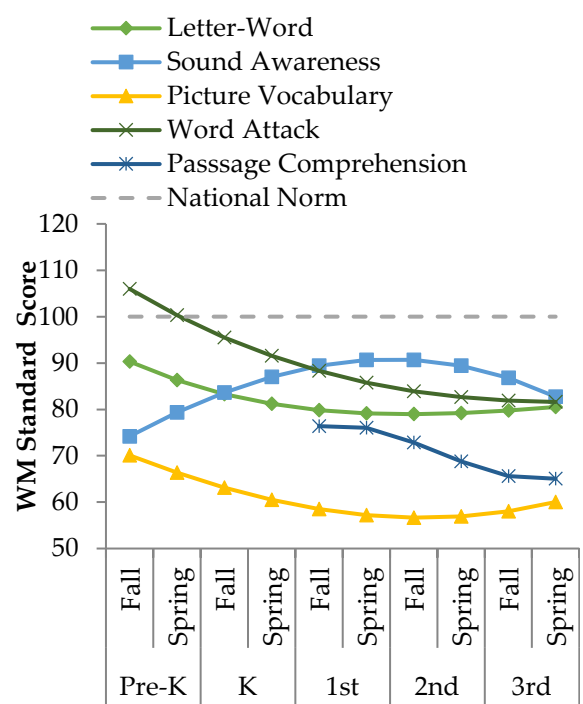


Figure 13. DLL Growth in Math Skills Assessed in English

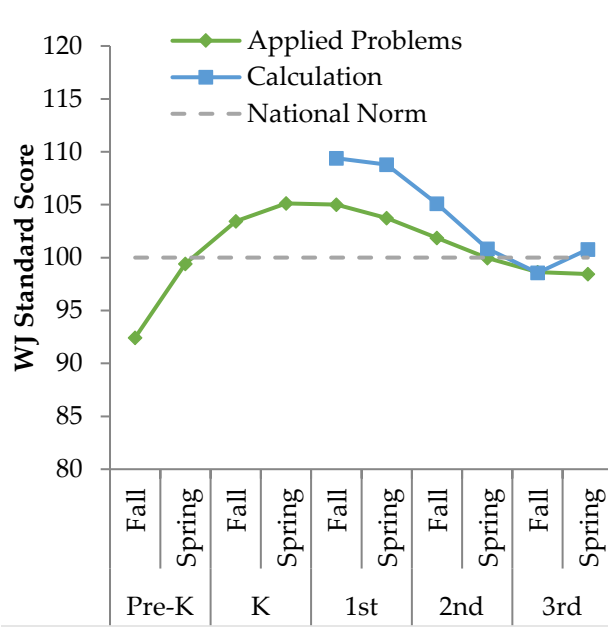


Figure 14. DLL Growth in Math Skills Assessed in Spanish

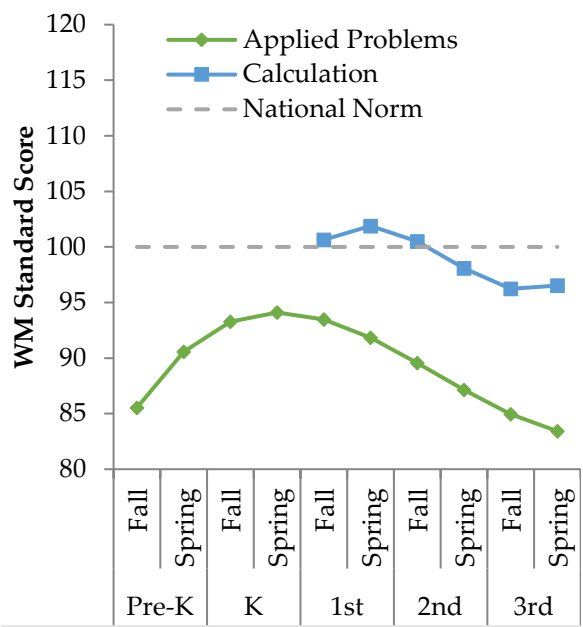
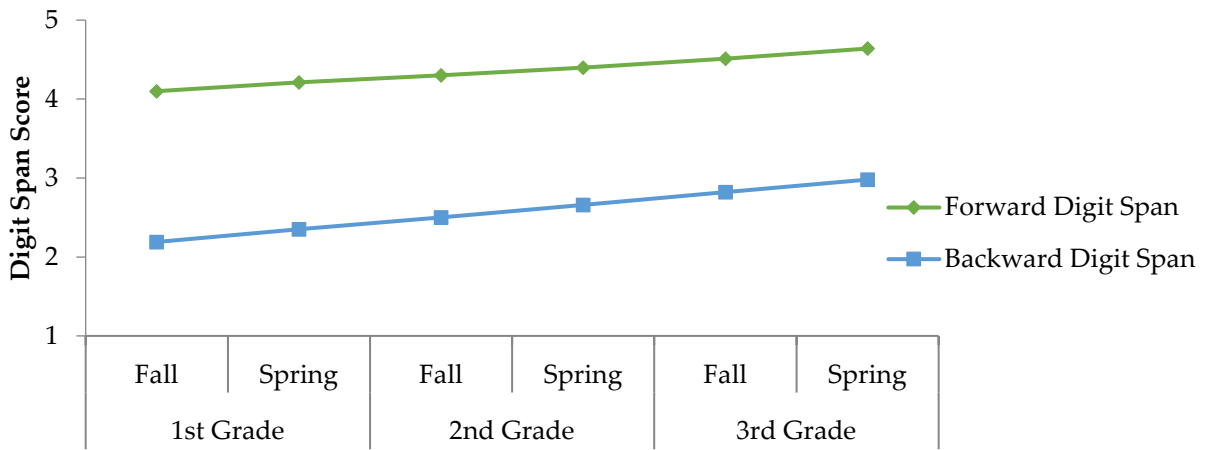


Figure 15. DLL Growth in Executive Function Skills*



*Figure 15 Digit Span measures produce raw scores, not standardized scores, so national averages are not available. These skills were only assessed in English.

Predictors of Patterns of Growth in Children's Scores Over Time

To examine predictors of growth in children's scores over time, moderation analyses examined if children's growth (time) in academic and behavior skills were similar for children with different English proficiency levels, child sex, having an individualized education plan (IEP status) at the beginning of pre-k, racial and ethnic backgrounds, and pre-k (public/private, total CLASS score) and K-3 classroom characteristics (high versus moderate/low quality classrooms).^a To adjust for multiple comparisons that increase the likelihood of falsely identifying a statistically significant result, adjustments were applied to p-values.^{iv} The following moderators predicted differences in growth on some longitudinal outcomes for children. The results from the moderation analyses were as follows:

- The following moderators were not associated with statistically different patterns of growth: children's racial and ethnic background and pre-k characteristics (pre-k located in a public/private setting and total CLASS Pre-K score).
- Moderators that predicted different patterns of growth over time for some outcomes included: child English proficiency level at pre-k, IEP status at the beginning of pre-k, sex of child, and K-3 classroom characteristics (high versus moderate/low quality classrooms).

English Proficiency Level

Children's English proficiency level at pre-k entry was associated with differences in longitudinal growth in receptive vocabulary skills (Picture Vocabulary subtest; see [Figure 16](#)), quantitative reasoning (Applied Problems subtest; see [Figure 17](#)), and social skills (Social Skills Rating; see [Figure 18](#)). There were five levels of English proficiency with level 1 being the least

^a HLM analyses were conducted with interaction terms included for time and moderators of interest.

proficient and level 5 being proficient. There were too few children that tested at levels 2 and 3 to analyze separately, so these two levels were combined for comparison purposes.

Figure 16. Growth in Picture Vocabulary Skills Over Time by Pre-K English Language Proficiency Level

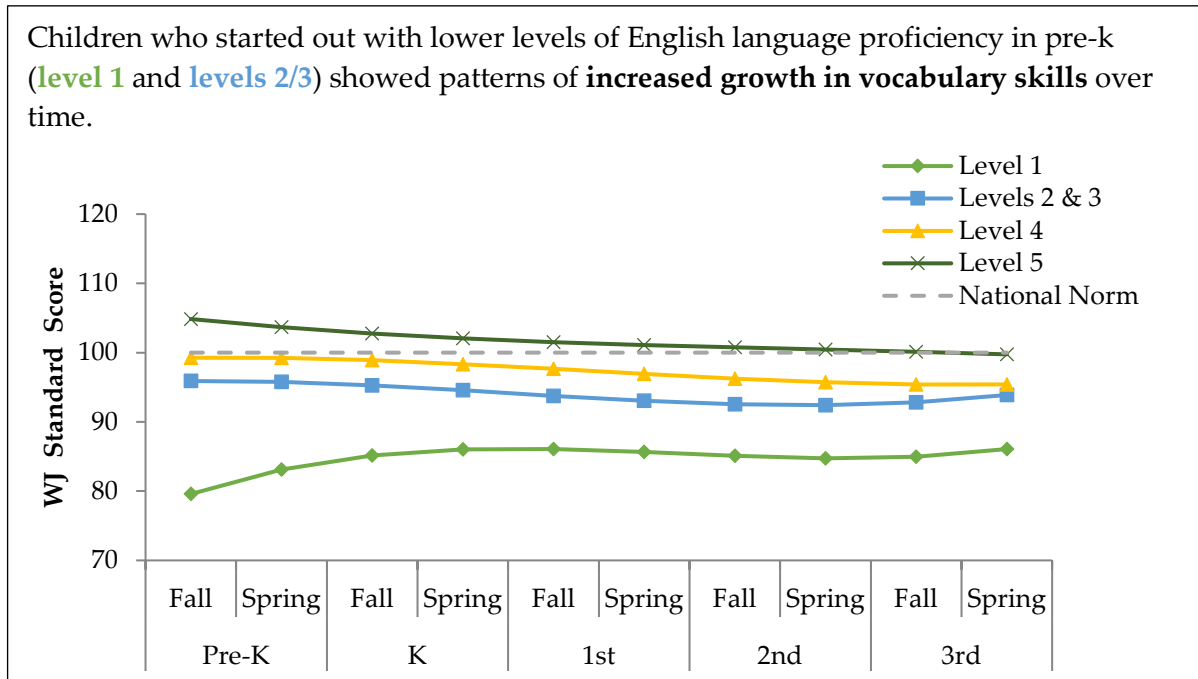


Figure 17. Growth in Applied Problems Skills Over Time by Pre-K English Language Proficiency Level

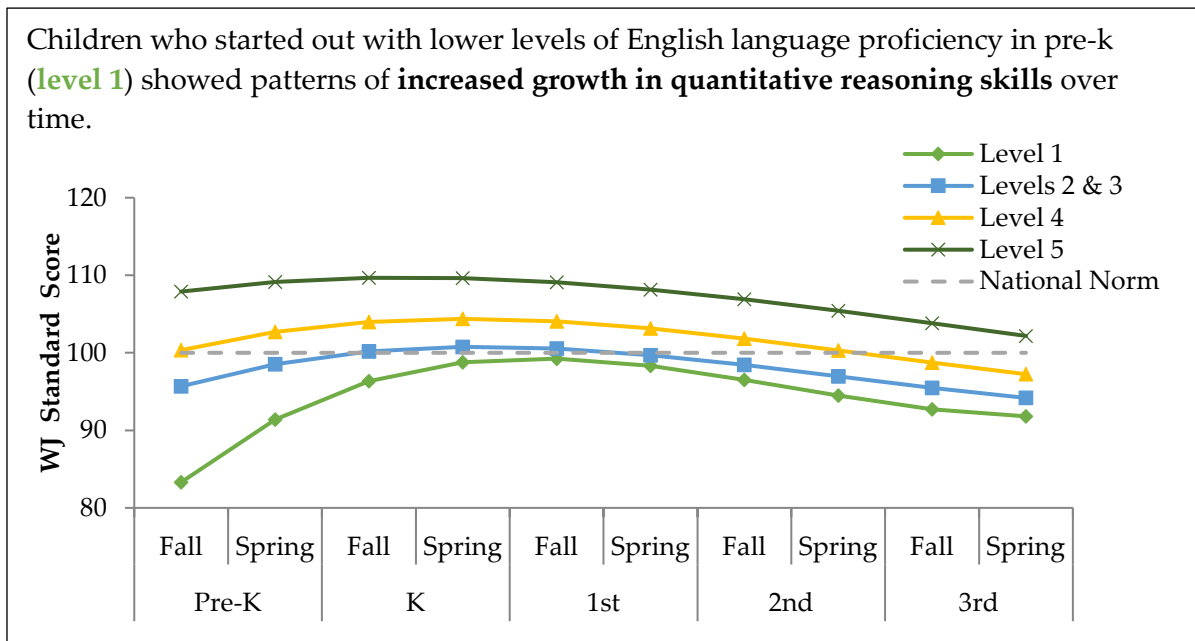
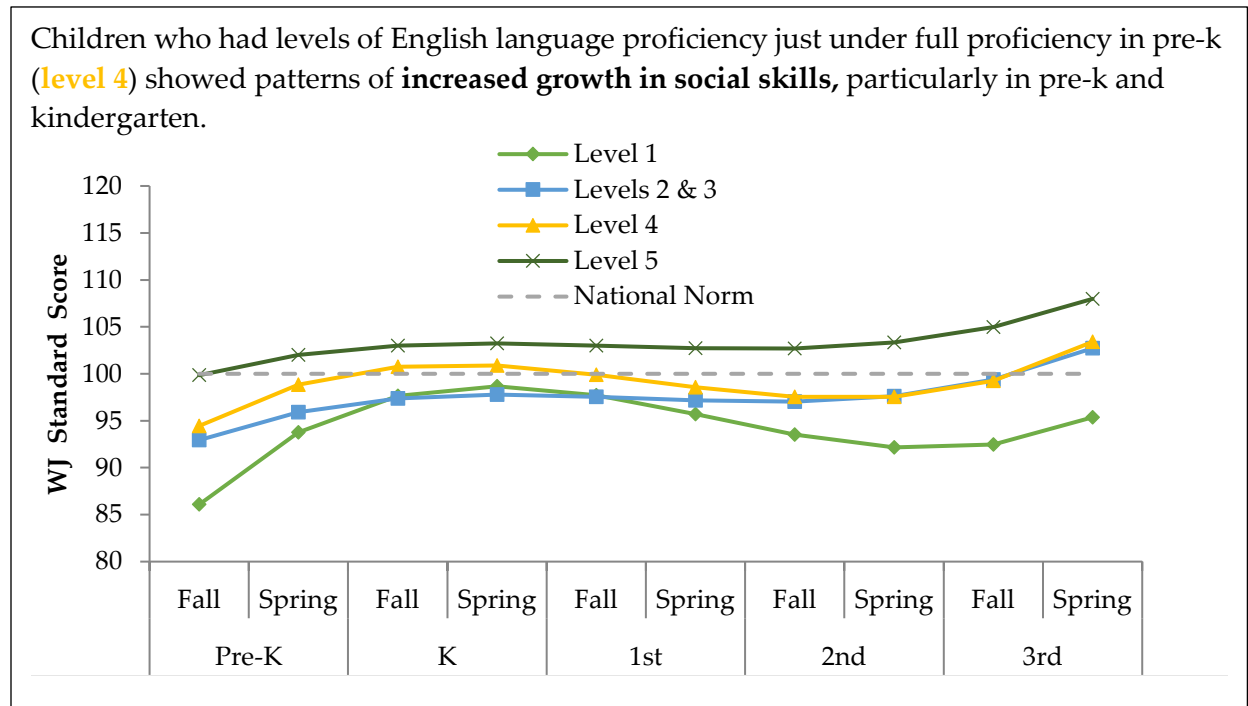


Figure 18. Growth in Social Skills Over Time by Pre-K English Language Proficiency Level



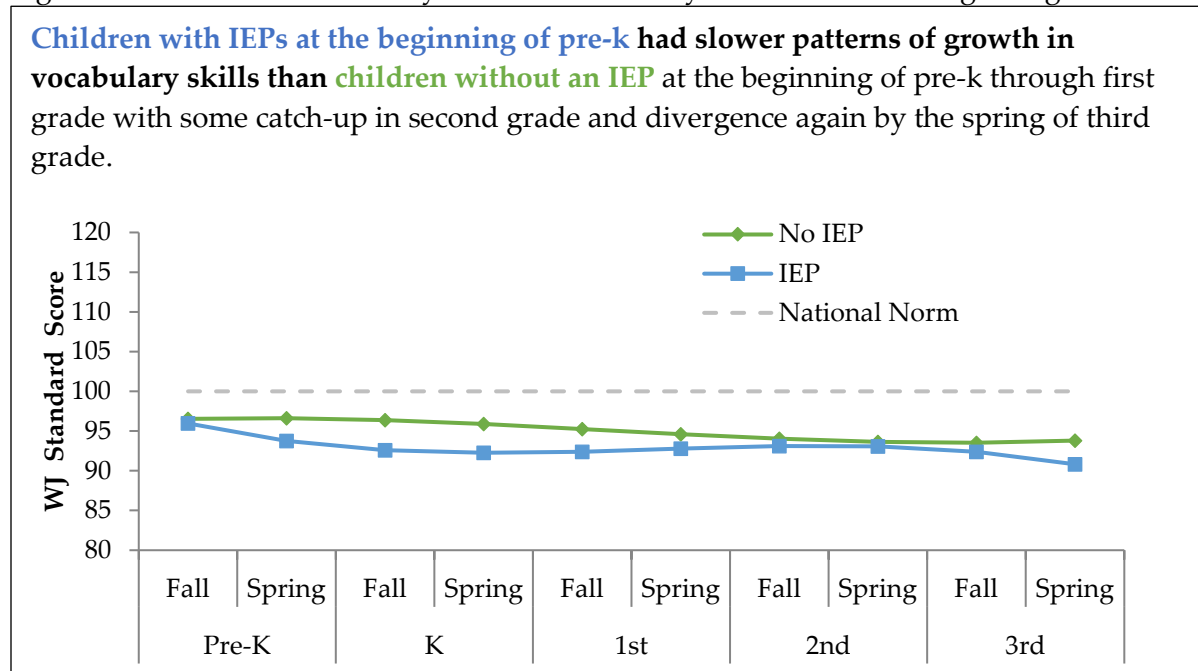
Children with IEPs at the Beginning of Pre-K and Sex of the Child

Two other child characteristics were significantly related to child outcomes over time.

Vocabulary skills did not improve as rapidly for children with IEPs at the beginning of pre-k (see [Figure 19](#)). However, children with an IEP at the beginning of pre-k did catch-up in second grade, but fell below peers again by the spring of third grade. The sex of the child was also significantly associated with patterns of change in problem behaviors over time. Girls had lower levels of problem behaviors in pre-k and third grade than boys with similar levels of problem behaviors in the spring of kindergarten through second grade (see [Figure 20](#)).

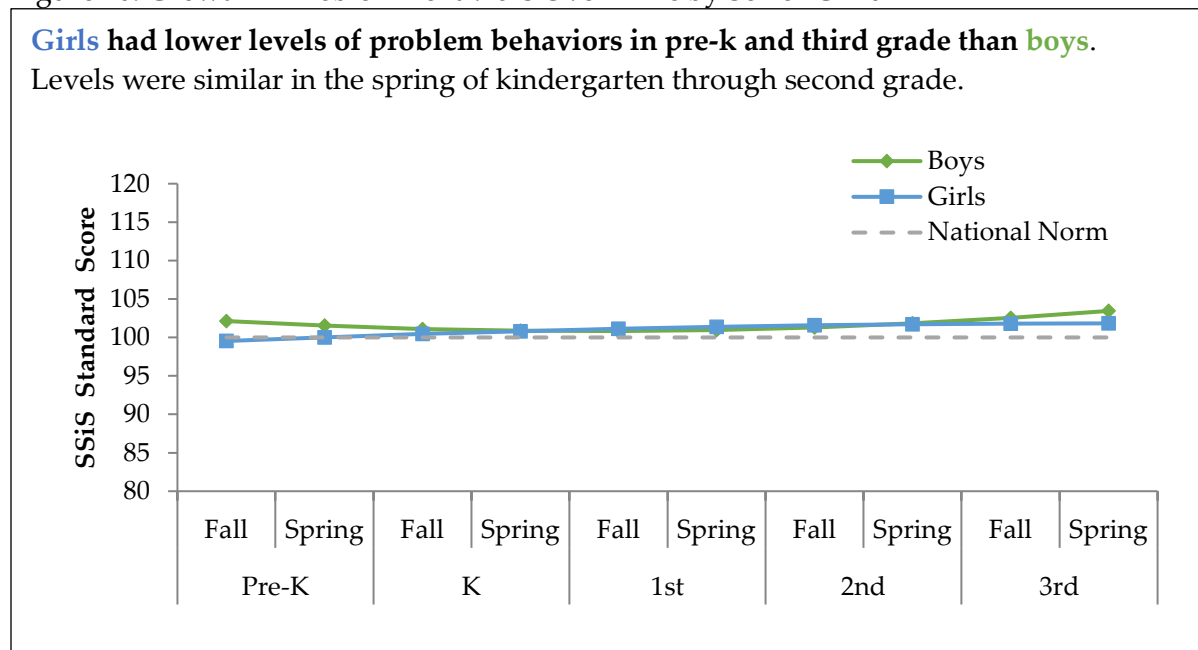
Children with IEPs at Beginning of Pre-K

Figure 19. Growth in Vocabulary Skills Over Time by IEP Status at the Beginning of Pre-K.



Sex of Child

Figure 20. Growth in Problem Behaviors Over Time by Sex of Child



Pre-K Program/Classroom Characteristics

Enrollment in high-quality classrooms compared to classrooms of moderate to low quality was associated with very small differences in longitudinal growth in phonemic awareness and letter identification (Letter-Word Identification subtest; see [Figure 21](#)). Children in high-quality classrooms in pre-k initially had better skills, but children enrolled in moderate-quality classrooms in kindergarten through third grade caught up as children who attended high-quality classrooms had scores closer to average over time. Similar patterns were observed for the Word Attack subtest (see [Figure 22](#)). Patterns of growth related to problem behaviors and classroom quality changed over time. From the spring of kindergarten through the spring of first grade, children in high-quality classrooms had fewer behavior problems and more behavior problems in third grade than children in low- or moderate-quality classrooms (see [Figure 23](#)).

Figure 21. Growth in Letter-Word Identification Skills Over Time by Kindergarten–Third Grade Classroom Quality: Low/Moderate versus High

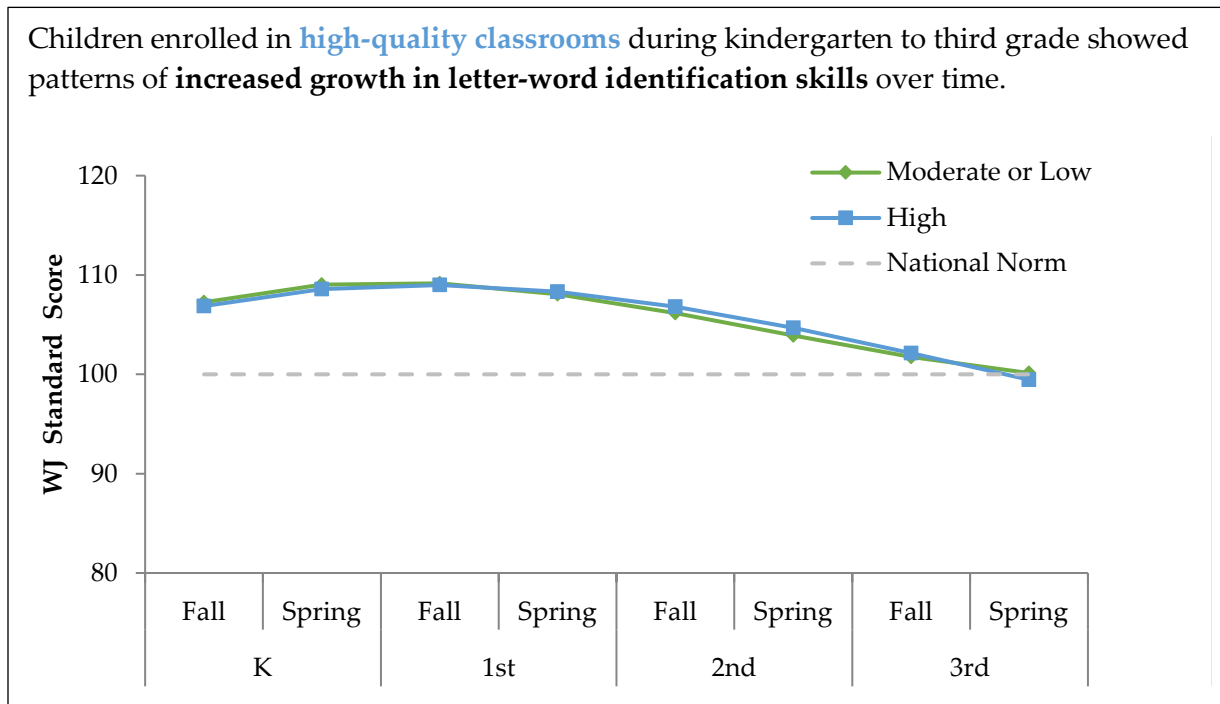


Figure 22. Change in Word Attack Subtest Scores Over Time by Kindergarten–Third Grade Classroom Quality: Low/Moderate versus High

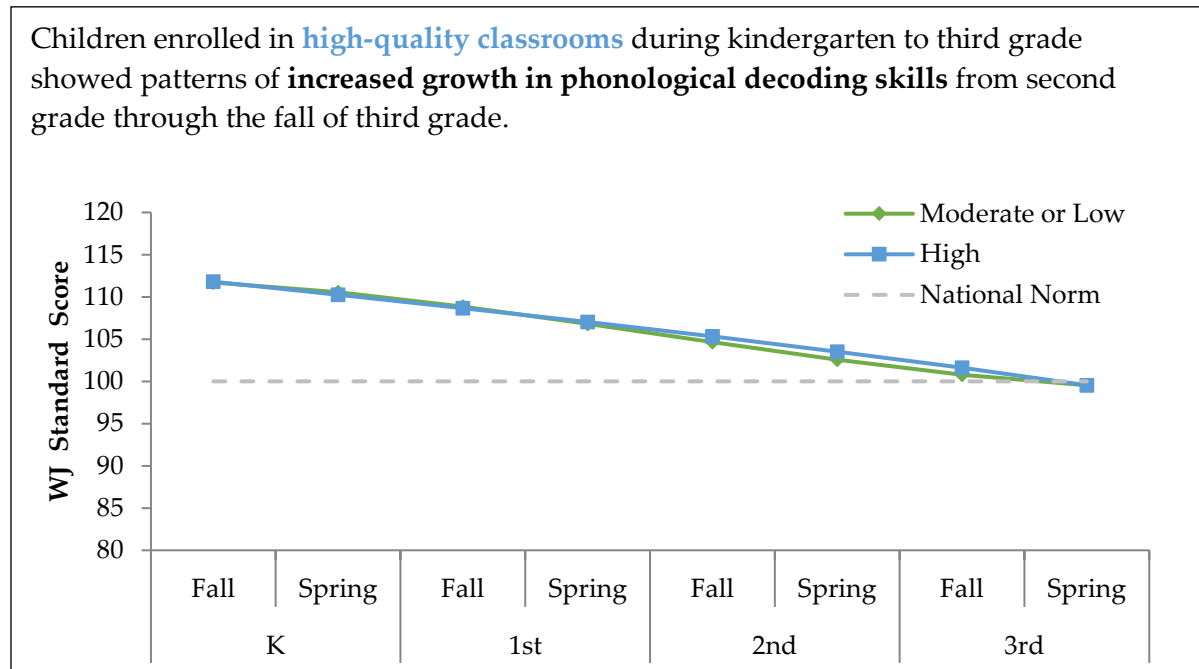
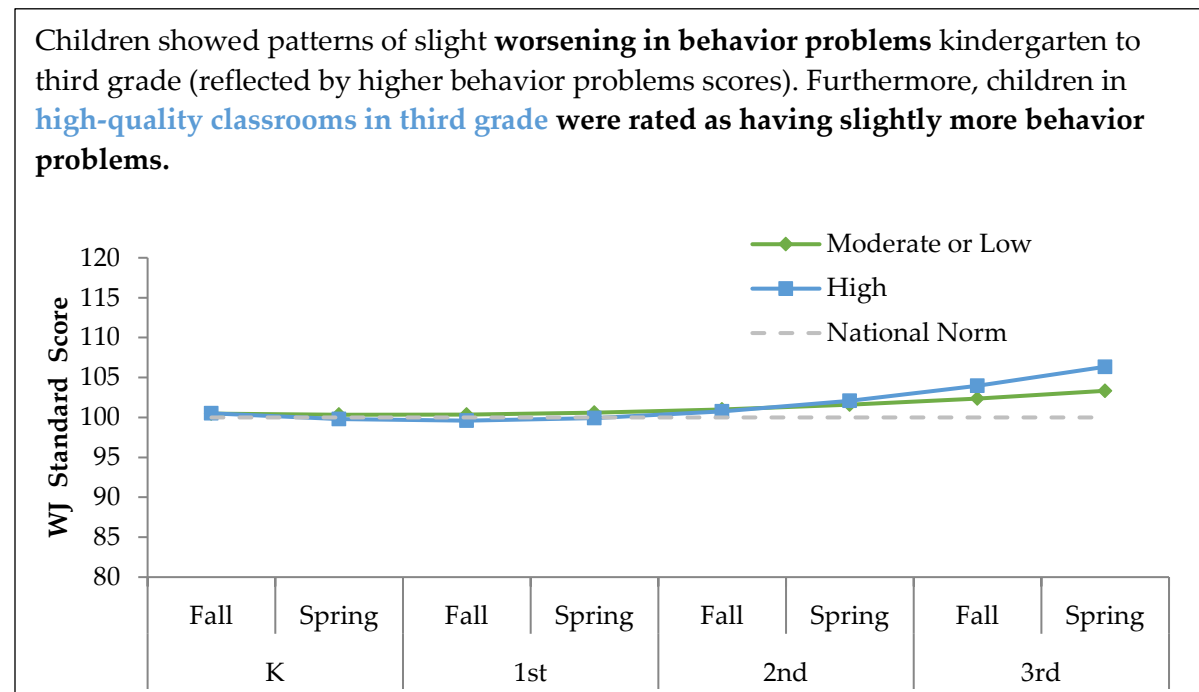


Figure 23. Growth in Problem Behaviors Over Time by Kindergarten–Third Grade Classroom Quality: Low/Moderate versus High



Classroom Quality

Researchers examined the quality of instructional practices each year from pre-k through third grade for the observed samples of classrooms attended by children in the study. They conducted descriptive analyses for CLASS Pre-K and kindergarten, first, second, and third grade CLASS K-3 scores including means and frequency distributions on the total, domain, and dimension scores.

Average classroom quality total scores were in the mid-range across all years (scores of 3.0-5.9; see [Figure 24](#)). Classroom quality was highest in the pre-k year and was slightly lower and consistent through third grade. Emotional Support and Classroom Organization domain scores were at the high end of the mid-range of quality, with Emotional Support scores increasing slightly from kindergarten through third grade. Classroom Organization scores were slightly lower in first and second grade and increased to the same levels as pre-k by third grade. Finally, similar to most other classrooms, Instructional Support scores were in the low end of the range (scores of 1.0-2.9) and decreased slightly from pre-k to second grade with a small increase in third grade (see [Figure 245](#)). A slightly higher proportion of programs were in the high-quality range in third grade compared to first and second grade for the total score (14%), Emotional Support (79%), and Classroom Organization domains (76%; see [Figures 25, 26, 27, and 28](#)). A slightly higher proportion of classrooms were in the mid-range for the Instructional Support domain in third grade (16%) compared to first and second grade.

Figure 24. Frequency of CLASS Average Total Scores by Grade

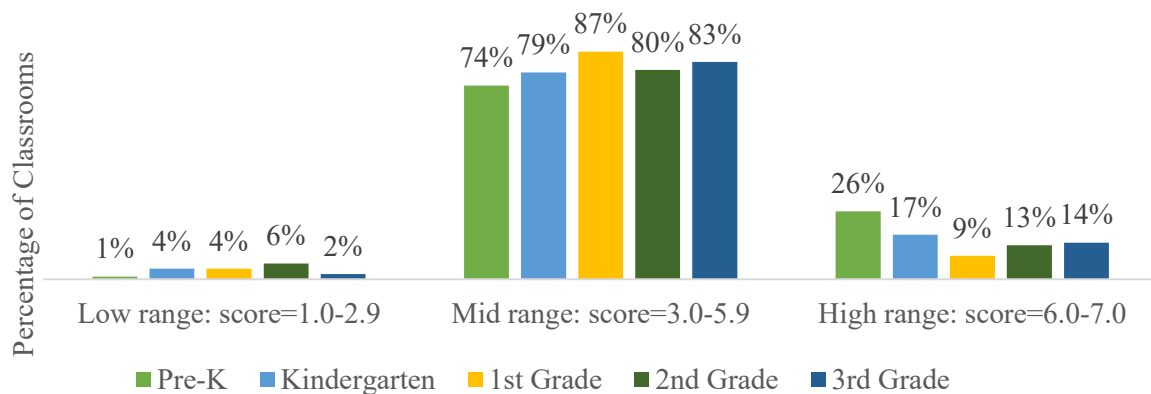


Figure 25. Average CLASS Total Scores by Grade

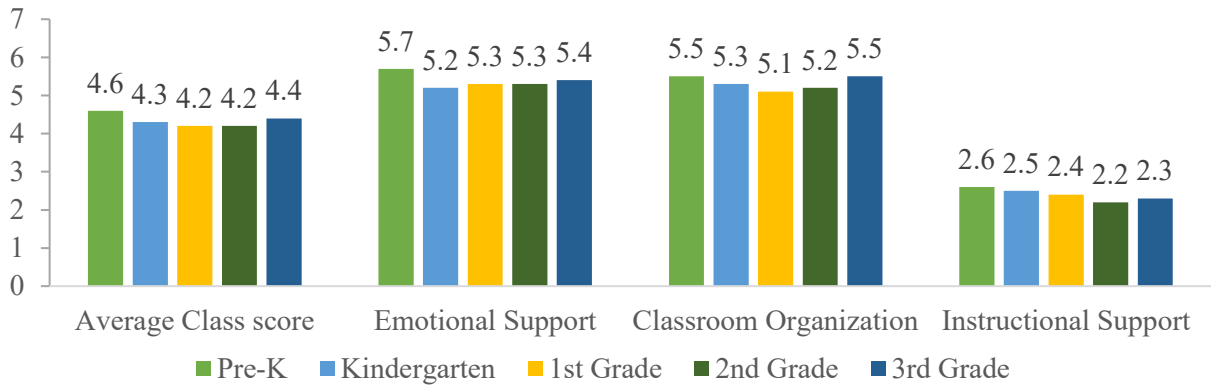


Figure 26. Frequency of Average CLASS Scores: Emotional Support Domain

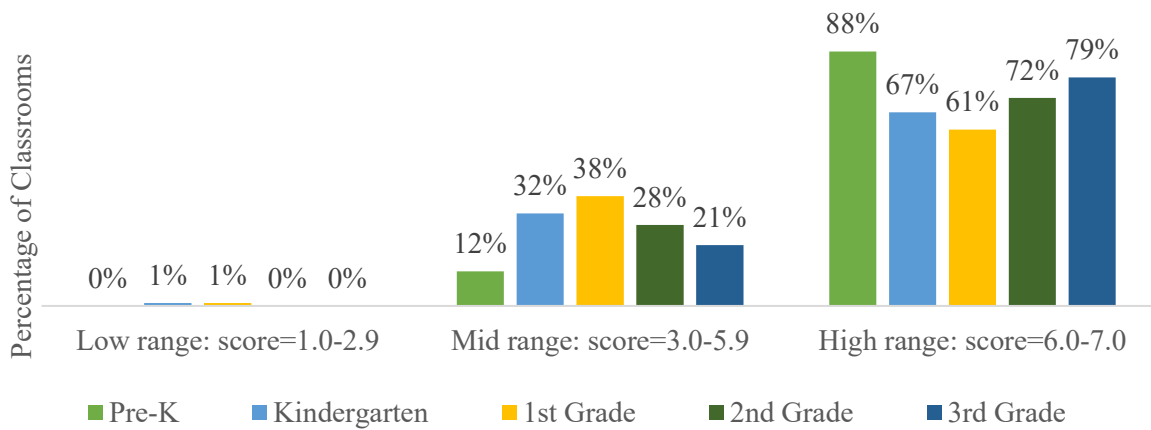


Figure 27. Frequency of Average CLASS Scores: Classroom Organization Domain

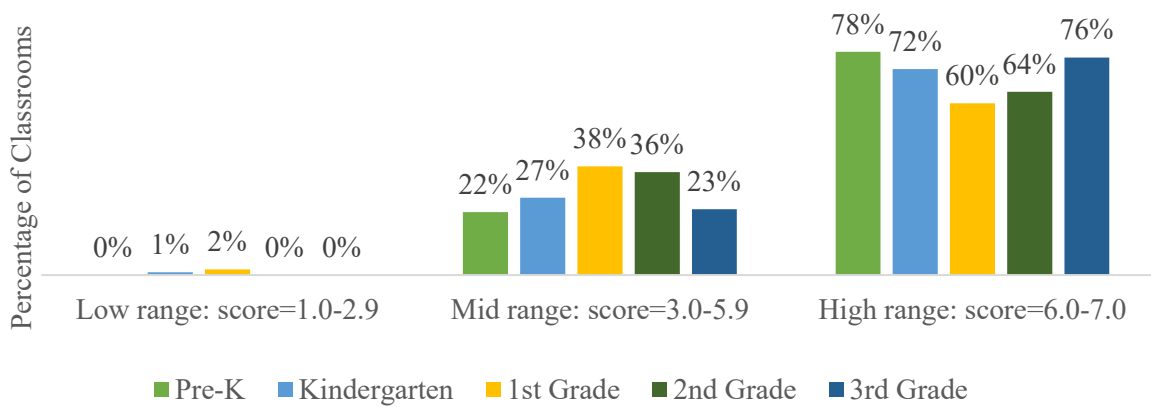
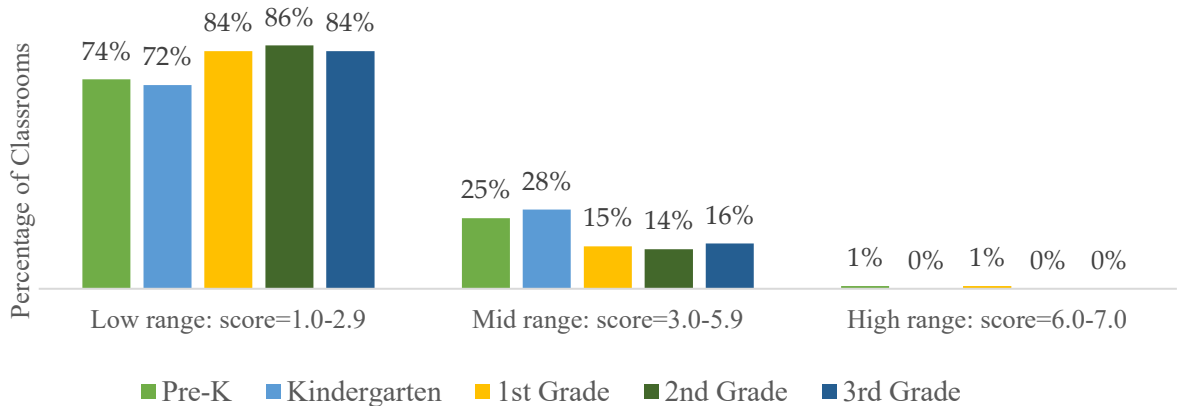


Figure 28. Frequency of Average CLASS Scores: Instructional Support Domain



Third Grade Longitudinal Study Summary

Pre-K through Third Grade within Grade Gains and Change Over Time

The following section highlights key findings from the longitudinal portion of the current study where children were assessed while enrolled in Georgia’s Pre-K Program through third grade:

- **Literacy:** Children who attended Georgia’s Pre-K Program continued to score slightly above the national average for reading skills related to decoding and phonics.
- **Vocabulary:** Children’s vocabulary skills started out within the average range relative to peers of the same age in pre-k and declined into the lower end of that range by third grade.
- **Math** skills were within the average range and were relatively stable over time for quantitative reasoning, math achievement, and math knowledge skills and at the higher end of the average range with declines to scores closer to the average score for children of the same age related to computational skills.
- **Executive function** skills grew steadily during the study period.
- **Dual language learners** showed a similar pattern of growth in English and Spanish; however, vocabulary scores were in the lower end of the average range in English and improved over time, and Spanish vocabulary skills were in the below-average range.

Comparing pre-k baseline scores to third grade scores, some pre-k gains were sustained through third grade, with children having higher-than-expected scores after controlling for child, family, and classroom characteristics in reading decoding (L-W ID), decoding and spelling (SA), and social skills. Very small to small decreases in the other norm-referenced outcomes were observed related to the remaining language and literacy and math skills. No norming sample was available for the executive function outcomes. These results are consistent with other longitudinal studies of pre-k effects.

Predictors of Longitudinal Outcomes

Most of the child and classroom-level moderators tested in this analysis were not associated with differential growth in child outcomes, including racial and ethnic backgrounds and pre-k characteristics—pre-k located in a public/private setting and total CLASS Pre-K score. However, children’s English proficiency level at pre-k was associated with growth in language and quantitative reasoning. Children who started out with the lowest levels of English language skills showed the greatest growth over time in these domains. Children with IEPs at the beginning of pre-k had slower patterns of growth in vocabulary skills than children without an IEP at the beginning of pre-k, and girls had slightly fewer problem behaviors at pre-k and third grade.

Classroom Quality Pre-K through Third Grade

Overall classroom quality was largely in the mid-range from pre-k through third grade. The proportion of classrooms that fell in the high-quality range increased slightly in third grade compared to grades one and two. Children enrolled in classrooms with high overall quality showed slightly accelerated growth in literacy skills and small increases in problem behaviors in pre-k and third grade compared to children enrolled in low- and mid-quality classrooms.

Comparison Sub-study

Comparison of Third Grade Scores: Georgia's Pre-K versus No Pre-K

The cohort of children included in the longitudinal study of Georgia's Pre-K Program was followed from pre-k through third grade. Tracking growth over time using standardized scores provides some understanding of how well children who attended Georgia's Pre-K Program perform relative to a norming sample of same-age peers. However, the longitudinal analyses did not provide insight into how children would have performed had they not attended Georgia's Pre-K Program. To examine if children who enrolled and did not enroll in Georgia's Pre-K Program have differential outcomes at third grade, a sub-study was conducted to examine differences in academic and social outcomes between the two groups in the fall of third grade.

A nested cohort design for the sub-study was used where the pre-k (treatment) group was composed of children who had enrolled in Georgia's Pre-K Program and had participated in the longitudinal study in the fall of third grade. The comparison group for this cohort was composed of 333 third grade children enrolled in the same third grade classes as the pre-k cohort children for the 2017-2018 school year but who had no experience with any pre-k during the 2013-2014 school year as reported by their parent or primary caregiver. Baseline assessments for the comparison group were administered in the fall of third grade for the beginning of this separate sub-study. Differences were estimated in academic and social skills for children with exposure to Georgia's Pre-K Program compared to children who had no pre-k exposure.

Comparison Group Recruitment

To identify a comparison group as similar as possible to the longitudinal sample, the recruitment process for the new no pre-k group included contacting parents from the same third grade classrooms attended by the longitudinal pre-k sample. Each classroom (n=667) was sent 25 recruitment packets to be sent home to parents. Parents who responded to the recruitment request completed a consent form and answered a short survey. From these 3,682 recruitment survey responses, most were not qualified due to having attended pre-k. Therefore, consent for 333 children (80 Spanish-speaking DLLs) was gathered from third graders who had no pre-k experience (birthdates ranging from 9/2/2008 to 9/1/2009).

Parent and Teacher Surveys

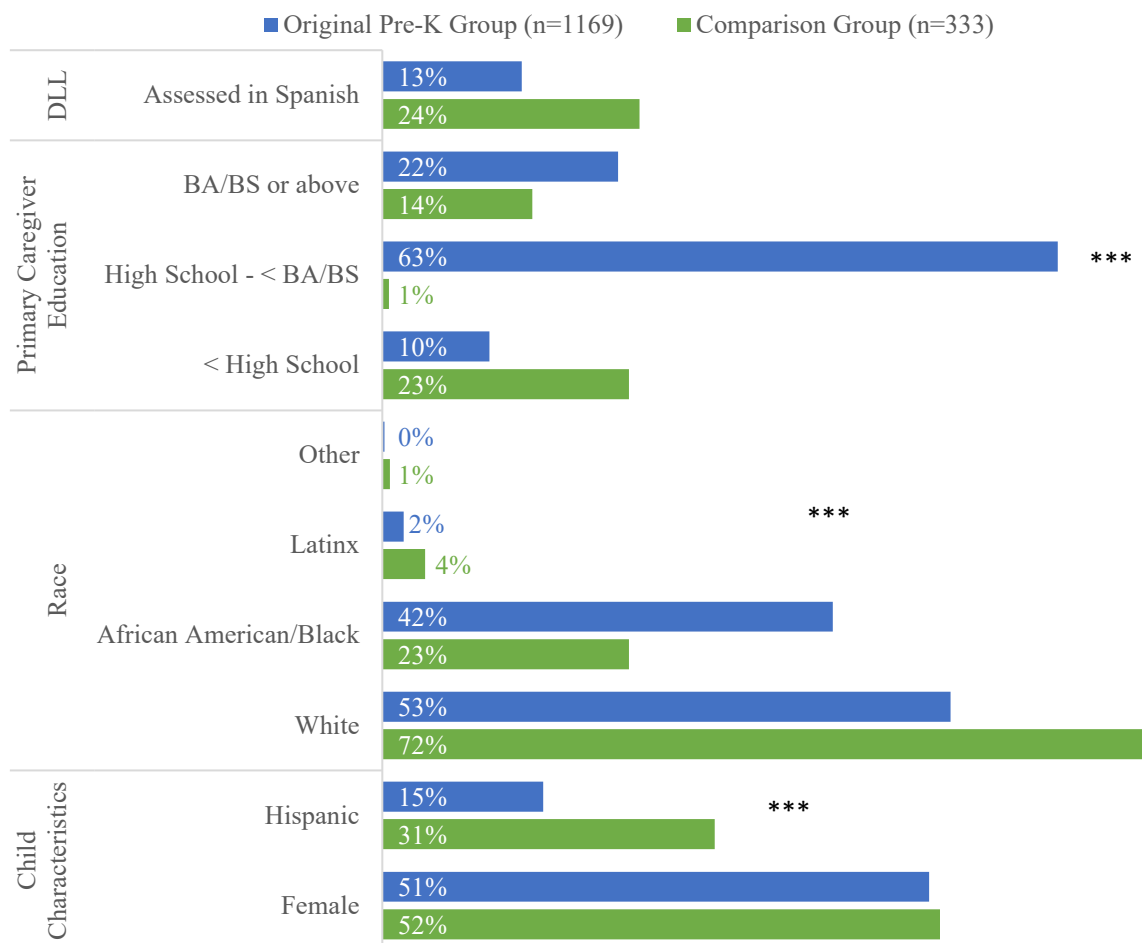
The third grade fall parent surveys included information about parent education levels, income, languages spoken, and number of adults/children in the home. To ensure that the children in the comparison group had not attended any pre-k, parents of children in the comparison group were asked if their child had attended pre-k/preschool as a four year old, and if so, the number of hours/day and months attended. Parent surveys were distributed to families through the classrooms and returned to teachers in sealed envelopes to be retrieved by the research team.

Parent surveys were received from 100% (333/333) of the participating third grade families whose children had not participated in Georgia’s Pre-K Program.

Comparison Group Participants

Information about child and family characteristics for the comparison group in third grade was obtained from parent survey data (see [Figure 29](#)). For the comparison group analyses, ethnicity (Hispanic or not) and race (Black, White, Latinx, and other) were treated separately. Results from t-test comparisons between the third grade longitudinal pre-k sample and the comparison group characteristics showed that the comparison group was more likely to be Hispanic (31% versus 15%), White (72% versus 52%) or Latinx (4% versus 2%); had primary caregivers with less education (52% with a high school diploma or below versus 43%); lower income (134% of the federal poverty level versus 163%); and had slightly larger household sizes (five people versus four).

Figure 29. Third Grade Comparison Group versus Original Pre-K Group



*Std Diff: Absolute Standardized Difference (Hodges-Lehmann Estimator).

* p-value < .05, ** p < .10, *** p < .001

Child Assessments

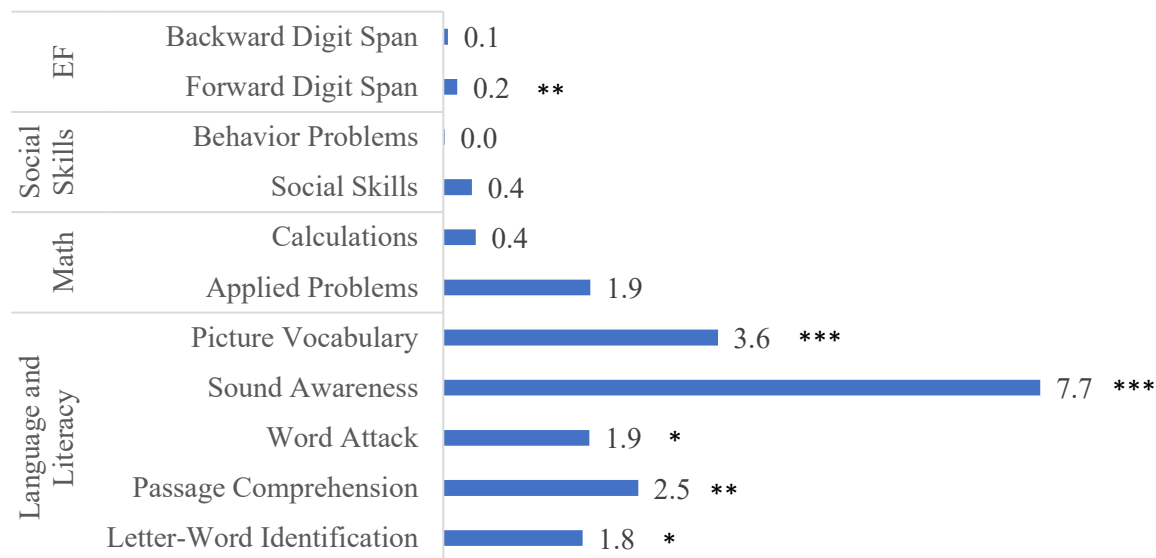
All children in the comparison group received the same assessments as the students in the longitudinal sample described above.

Comparison Group Analyses

To estimate differences in academic and social outcomes at the fall of third grade for children who had been enrolled in Georgia’s Pre-K Program compared to children who were not enrolled in any pre-k program, separate generalized linear models were fit for each outcome. These models estimated differences in academic and social outcomes at the fall of third grade between the group who attended Georgia’s Pre-K Program and the group who did not attend any pre-k, as an approximation of the effect of retrospective pre-k exposure. These models controlled for the demographic characteristics on which the comparison group differed (see [Figure 29](#)).

Children who attended Georgia’s Pre-K Program had significantly better scores on all language and literacy outcomes in the fall of third grade compared to children who did not attend any pre-k program. In particular, children who attended Georgia’s Pre-K Program benefitted most in Sound Awareness compared to their peers who did not attend pre-k (see [Figure 30](#)). Differences were statistically significant for Executive Function skills as measured by the Forward Digit Span task. Differences in language and literacy outcomes were in the small to moderate range (Effect Size = 0.12-0.51).

Figure 30. Differences in Standard and Raw Scores for Pre-K Group versus Comparison Group in Third Grade



* p-value < .05, ** p < .10, *** p < .001

Summary and Conclusions on the Comparison Sub-study

Children who attended Georgia's Pre-K Program had significantly better scores than their peers who did not attend any pre-k program on the following outcomes:

- Literacy (ordered from least to most complex skills):
 - Letters, letter sounds, phonological, and orthographic coding (Letter-Word Identification subtest)
 - Phonological/orthographic coding with nonsense words (Word Attack)
 - Reading comprehension and lexical knowledge (Passage Comprehension subtest)
- Oral Language (ordered from least to most complex):
 - Rhyming and manipulating phonemes (Sound Awareness subtest)
 - Identifying and naming pictures (Picture Vocabulary subtest)
- Executive Function:
 - Attention efficiency and capacity (Forward Digit Span Task)

The largest difference between the comparison and pre-k groups was related to the Sound Awareness subtest. This subtest assesses auditory processing skills that underlie decoding and spelling as basic skills for acquiring literacy. In the sequence of skills necessary for literacy acquisition, these skills are some of the least complex and are most likely to be targeted with young children in a pre-k program. The difference between groups was of moderate size in third grade (Effect Size = 0.51). This effect size is similar, but slightly smaller, than the fall-to-spring gain in pre-k for the longitudinal study group (Effect Size = 0.61). If these skills are not acquired early in literacy acquisition, they are unlikely to be taught later.

Appendix A. Child Outcome and Classroom Quality Measures

Measure	Scoring
Language and Literacy Skills	
Vocabulary WJ-III ^v Picture Vocabulary (Subtest 14) / Bat-III ^{vi} Vocabulario sobre Dibujos (Prueba 14)	Standard score Mean=100, SD=15
Phonological Awareness WJ-III Sound Awareness (Subtest 21) / Bat-III Discernimiento de Sonidos (Prueba 21)	Standard score Mean=100, SD=15
Letter-and Word Recognition WJ-III Letter-Word Identification (Subtest 1) / Bat-III Identificación de Letras y Palabras (Prueba 1)	Standard score Mean=100, SD=15
Written Comprehension WJ-III Passage Comprehension (Subtest 9) / Bat-III Comprension de Textos (Prueba 9)	Standard score Mean=100, SD=15
Phonemic Awareness and Decoding WJ-III Word Attack (Subtest 13) / Bat-III Análisis de Palabras (Prueba 13)	Standard score Mean=100, SD=15
Math Skills	
Written Calculation WJ-III Calculation (Subtest 5) / Bat-III Cálculo (Prueba 5)	Standard score Mean=100, SD=15
Math Problem-Solving WJ-III Applied Problems (Subtest 10) / Bat-III Problemas Aplicados (Prueba 10)	Standard score Mean=100, SD=15
Executive Function	
Working Memory Forward Digit Span and Backward Digit Span (English/Spanish) ^{vii}	Raw score Range=1-8
Behavior Skills	
Social Skills SSiS ^{viii} Social Skills subscale	Standard score Mean=100, SD=15
Problem Behaviors SSiS Problem Behaviors subscale	Standard score Mean=100, SD=15
Classroom Quality	
Teacher-Child Instructional Interactions CLASS ^{ix} / CLASS K-3 ^x Emotional Support, Classroom Observation, Instructional Support Domains & Total	Total and Domain scores range=1.0–7.0

References

-
- ⁱ Barnett, S.W., Jung, K., Youn, M., & Frede, G.C. (2015). Abbott Preschool Program Longitudinal Effects Study: Fifth Grade Follow-Up. *National Institute for Early Education Research at Rutgers-The State University of New Jersey*. Retrieved October 1, 2020, from www.nieer.org
- ⁱⁱ Hill, C.J., Gormley, W.T., & Adelstein, S. (2015). Do the short-term effects of a high-quality preschool program persist? *Early Childhood Research Quarterly*, 32, 60-79.
- ⁱⁱⁱ Lipsey, M.W., Farran, D.C., & Durkin, K. (2018). Effects of the Tennessee Prekindergarten Program on children's achievement and behavior through third grade. *Early Childhood Research Quarterly*, 45, 155-176. <https://doi.org/10.1016/j.ecresq.2018.03.005>
- ^{iv} Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)*, 57(1), 289-300.
- ^v Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III Tests of Achievement*. Itasca, IL: The Riverside Publishing Company.
- ^{vi} Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III Tests of Achievement*. Itasca, IL: The Riverside Publishing Company.
- ^{vii} Gathercole, S. E., & Pickering, S. J. (2000). Working memory deficits in children with low achievements in the national curriculum at 7 years of age. *British Journal of Education Psychology*, 70, 177-194.
- ^{viii} Gresham, F., & Elliott, S. (2008). *Social Skills Improvement System*. Minneapolis, MN: Pearson.
- ^{ix} Pianta, R. C., La Paro, K. M., Hamre, B.K. (2008) *Classroom Assessment Scoring System (CLASS) Pre-K Manual*. Baltimore, MD: Paul H. Brookes Publishing Co.
- ^x Pianta, R. C., La Paro, K. M., Hamre, B.K. (2008) *Classroom Assessment Scoring System (CLASS) K-3 Manual*. Baltimore, MD: Paul H. Brookes Publishing Co.