Georgia's Pre-K Program Evaluation Project

Children's Outcomes through Second Grade

Findings from Year 4 of Georgia's Pre-K Longitudinal Study

UNC
FRANK PORTER GRAHAM
CHILD DEVELOPMENT INSTITUTE

Ellen Peisner-Feinberg, PhD
Karen Van Manen, PhD
Irina Mokrova, PhD &
Margaret Burchinal, PhD

We wish to acknowledge the members of our Georgia's Pre-K Program Evaluation Team who assisted with the 2016-2017 study: Research Staff 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, Karen Stewart, Gina Walker; and Data Collectors Jane Argo, Joan Boudousquie, Elizabeth Bragg, Brianna Castellano, Cynthia Coleman, Abby Davis, CL Dunn, Deborah Hallman, Stacey Harris, Melissa Johns, Deborah Laing, Elizabeth Milling, Ann O'Mahoney, Jacqueline Parker, Nicole Poucher, Katie Rutland, Carolina Sandoval, Melissa Silva, Patricia Simpson, Moneesha Smith, Alissa Sypsa, Anna Tussey, Peggy Tymes, Adriana Vivas-Sosa, Ann Zimmer-Shepherd.

In addition, we offer 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 the Georgia's Pre-K Program and the kindergarten, first-grade, and second-grade classrooms and schools; and the staff of Bright from the Start: Georgia Department of Early Care and Learning.

Cover by Gina Harrison.

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: Peisner-Feinberg, E., Van Manen, K., Mokrova, I., & Burchinal, M. (2019). *Children's Outcomes through Second Grade: Findings from Year 4 of Georgia's Pre-K Longitudinal Study.* 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 www.decal.ga.gov.

Contents

List of Tables

List of Figures

Purpose of the Evaluation Study

In 2011-2012, the Georgia legislature funded a series of ongoing studies to evaluate Georgia's Pre-K Program. The first study, conducted in 2011-2012, was designed to examine children's learning outcomes during pre-k, the factors that predict better outcomes, and the quality of children's experiences in Georgia's Pre-K classrooms based on a random sample of 100 classrooms and 509 children within those classrooms. The second study, conducted in 2012-2013, was designed to investigate the effects of participation in Georgia's Pre-K on children's school readiness skills, and whether those effects are similar for different groups of children. This study utilized a regression discontinuity design (RDD) to compare children who had and had not attended the program, and included 1,181 children (611 treated and 570 untreated). The third study, occurring from 2013-2020, involves a longitudinal design to follow a sample of 1,169 children from pre-k through fifth grade, in order to examine the short- and long-term learning outcomes for children who attended Georgia's Pre-K as well as the quality of their preschool and school experiences.

The 2016–2017 Georgia's Pre-K Program Evaluation focuses on the results of the fourth year of this longitudinal study, through second grade. The purpose of this evaluation study was to examine longitudinal outcomes for children related to key academic and social skills as well as the quality of their classrooms from pre-k through second grade. The primary evaluation questions included: 1) What are the learning outcomes through second grade for children who attended Georgia's Pre-K Program?, 2) What factors predict better learning outcomes for children?, and 3) What is the quality of children's instructional experiences from pre-k through second grade?

To address these questions, the evaluation study included a sample of 1,169 children (139 Spanish-speaking dual language learners/DLLs) attending a random sample of 199 Georgia's Pre-K classrooms in year 1. Of the original sample of children, 1,034 were followed into kindergarten in year 2 (118 Spanish-speaking DLLs), 969 (119 Spanish-speaking DLLs) were followed into first grade in year 3, and 951 (116 Spanish-speaking DLLs) were followed into second grade in year 4. Researchers conducted individual child assessments near the beginning and end of each year to examine growth in children's skills. The assessment measures covered multiple domains of learning, including language, literacy, math, executive function, and teacher ratings of behavior skills. For the DLL subsample, parallel assessments were conducted in both English and Spanish. Researchers also conducted observations in children's classrooms each year from pre-k through second grade to gather data on the quality of teacher-child instructional interactions using the CLASS. In addition, information about characteristics of the classrooms, teachers, and children was gathered from teacher and parent surveys and from existing statewide pre-k program data. Child/family characteristics, classroom/teacher characteristics, and classroom quality were examined as predictors of children's growth in skills.

Overview of Georgia's Pre-K Program

Georgia's Pre-K Program is a state-funded universal pre-kindergarten program for 4-year-olds. The program serves children from all income levels, with no fees charged to families for program participation. Georgia's Pre-K Program was established in 1992 and became one of the first states to offer a universal program in 1995. The program serves over 80,000 children each year in a variety of settings across the state, including public school systems, private providers, and blended Head Start/pre-k classrooms. Georgia's Pre-K is based on a school-year model with instruction for 180 days/year and 6.5 hours/day[a]. Class sizes are limited to 20–22 children with a lead and assistant teacher, and adult:child ratios of 1:11. Lead teachers are 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 are required to have at least a Paraprofessional Certificate (issued by the Georgia Professional Standards Commission) or a Child Development Associate (CDA) credential. In addition, program guidelines provide minimum salary requirements for lead teachers based on credentials, with funding provided by the state, as well as minimum salary requirements for assistant teachers meeting the credential requirements.

Guidelines for classroom instruction are provided through the *Georgia Early Learning and Development Standards (GELDS)*[i], which are aligned with *Georgia's Performance Standards for Kindergarten*[ii]. The program standards also require Georgia's Pre-K sites to use an approved curriculum; provide written lesson plans which include educational experiences in language and literacy, math, science, social studies, creative arts (music, art, and drama), social and emotional, and health and physical development; implement individual child assessments using the *Georgia's Pre-K Child Assessment—Work Sampling Online*[iii], which is based on the *Work Sampling System*[iv]; offer meals, rest time, and both indoor and outdoor play time; and provide support services or referrals to families as needed. Bright from the Start: Georgia Department of Early Care and Learning (DECAL) oversees the program, and staff provide consultation, technical assistance, and monitoring visits throughout the year. (See 2017–2018 Georgia's Pre-K Program Operating Guidelines[v] for further information.)

[a] Prior to 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.

Method

The current report includes data based on following a longitudinal sample of children from Georgia's Pre-K through second grade (see Table 1). In year 1 (2013-2014) of the longitudinal study, data were gathered from a random sample of classrooms and children within classrooms to examine child outcomes and classroom quality in Georgia's Pre-K Program. This random sample of children was then followed into kindergarten during year 2 (2014-2015), first grade during year 3 (2015-2016), and second grade during year 4 (2016-2017). At the beginning (fall) and end (spring) of each year, researchers conducted individual assessments of children's language and academic skills and gathered teacher ratings of behavior skills. Researchers also conducted classroom observations of teacher-child instructional interactions each year. Program and classroom characteristics, as well as teacher and child demographic data, were obtained from annual teacher and parent surveys and existing statewide administrative data collected by DECAL.

Participants

Classrooms

Children in the study sample initially attended 199 Georgia's Pre-K classrooms in year 1 (2013–2014), 822 kindergarten classrooms in year 2 (2014–2015), 777 first-grade classrooms in year 3 (2015-2016), and 786 second-grade classrooms in year 4 (2016-2017). (See Table 1.) About half of the Georgia's Pre-K Programs attended by children in the study sample were in public school settings (49%) and about half were in private sites (51%). In contrast, the elementary school classrooms attended by children in the study primarily were located in public school settings each year (K=97%, 1st=98%, 2nd=98%), with a few in charter schools (K=2%, 1st=1%, 2nd=1%) and private schools (K=1%, 1st=1%, 2nd=1%).

Information on characteristics of the classrooms and teachers included in the sample each year based on teacher surveys, as well as state administrative data from DECAL in pre-k, is presented in Table 2. The average class size was similar each year, about 20-21 children (PK=21, K=20, 1st=21, 2nd=21), with half boys and half girls. 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 the elementary grades, this pattern was reversed, with slightly over one-third having a bachelor's degree (K=38%, 1st=35%, 2nd=38%) and almost two-thirds having a master's degree or higher (K=62%, 1st=64%, 2nd=62%). The majority (82%) of pre-k teachers were Georgia PSC Certified or Certified Temporary. Teachers reported a substantial number of years of teaching experience on average (PK=11, K=15, 1st=14, 2nd=14), including around half of this time teaching their current grade level (PK=6, K=9, 1st=7, 2nd=6). The majority of teachers were White (PK=67%, K=77%, 1st=70%, 2nd=72%), almost one-quarter were Black (PK=27%, K=18%, 1st=23%, 2nd=23%); a small proportion were of Hispanic/Latino ethnicity (PK=3%, K=2%, 1st=2%, 2nd=2%), and nearly all were female (PK=98%, K=98%, 1st=99%, 2nd=97%).

Children

The study sample included 1,169 children in year 1 (pre-k), 1,034 children in year 2 (88% of the original sample in kindergarten), 969 children in year 3 (83% of the original sample in first grade), and 951 children in year 4 (81% of the original sample in second grade). These children included a subsample of Spanish-speaking dual-language learners (DLL subsample)—139 children in year 1 (pre-k), 118 children in year 2 (85% of the original sample in kindergarten), 119 children in year 3 (86% of the original sample in first grade), and 116 children in year 4 (83% of the original sample in second grade). Parent permission forms were distributed to all children in each randomly-selected pre-k classroom, with an overall permission rate of 73% (3,136 of 4,270 eligible children). An average of 6 children with parent permission per classroom were randomly selected for inclusion in the study in year 1. Children were excluded from year 2 of the study for the following reasons: parent withdrew permission (n=4), child attended an ineligible site (n=6), the research team was not able to locate the child in a school during kindergarten (n=30), child had moved out of state (n=33), or the school district or school was unwilling to participate in the study (n=62). Children were excluded from year 3 of the study for the following reasons: parent withdrew permission (n=19), child attended an ineligible site (n=3), the child was held back/retained during pre-k (n=6) or kindergarten (n=32), behavioral issues prevented assessment (n=1), the research team was not able to locate the child in a school during first grade (n=47), child had moved out of state (n=59), or the school district or school was unwilling to participate in the study (n=33). Children were excluded from year 4 of the study for the following reasons: parent withdrew permission (n=28); child attended an ineligible site (n=7); the child was held back/retained during pre-k (n=6), kindergarten (n=34), or first grade (n=16); the research team was not able to locate the child in a school during second grade (n=42); child had moved out of state (n=71); or the school district or school was unwilling to participate in the study (n=14).

Information about child and family characteristics for the study sample was obtained from DECAL data and parent survey data (see Table 3). The children in the sample in year 4 were about half boys (49%) and half girls (51%); from varied racial/ethnic backgrounds, including close to 40% each White (42%) and Black/African-American (37%), 15% Hispanic/Latino, and the remainder from other or multiracial backgrounds (6%). Approximately 10% of the children had limited English language proficiency; and 3% had an individualized education program (IEP). Slightly over half (54%) 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). The education level for the majority (62%) of children's primary caregivers was between a high school diploma and less than a bachelor's degree, with slightly over one-quarter (28%) having a bachelor's degree or above. Results from t-test comparisons between those included and not included in the year 4 sample (based on the original sample from year 1) found no significant differences between the two groups for any of these child or family characteristics.

Measures and Procedures

Child Assessments

Individual assessments to measure children's growth in skills were conducted in their pre-k, kindergarten, first-grade, and second-grade settings. Children were assessed at eight time points, in the fall and spring near the beginning and end of each year from pre-k through second grade. Assessments were conducted by data collectors trained by the research team, and children's verbal assent was obtained prior to the assessment. All children received assessments in English. Children who were reported by their parents or teachers to speak Spanish received a second set of parallel assessments using Spanish language versions of these measures. The Spanish assessments were conducted by a different, bilingual data collector on a separate day, approximately two weeks after the English assessments.

A battery of measures of language, literacy, and math skills was gathered longitudinally from pre-k through second grade. In first grade, more developmentally advanced measures of reading comprehension, number operations, and executive function were added to the battery. All of the child assessment measures were available in both English and Spanish versions. Most of the measures used were norm-referenced, and therefore, standard scores could be used for these measures. These scores take into account children's age, so that the standardized mean score of 100 represents the expected performance for an average child at a given age. (See Table 4 for a list of measures used in the study.)

Language and literacy skills were assessed with five subtests from the *Woodcock-Johnson III Tests of Achievement*[vi] (WJ-III) / *Batería III Woodcock-Muñoz Pruebas de Aprovechamiento*[vii] (Bat-III). The Picture Vocabulary subtest measured vocabulary skills, including aspects of both receptive and expressive language. The Sound Awareness subtest measured phonological awareness skills, including rhyming and ability to delete, substitute, and reverse sounds or word parts. The Letter-Word Identification subtest measured basic pre-reading and reading skills, including letter and word recognition and identification skills. The Word Attack subtest measured phonemic awareness and decoding skills, including knowledge of letter sounds and sound combinations. The Passage Comprehension subtest measured symbolic learning and basic written comprehension skills. Passage Comprehension was added in first grade; all other measures were gathered beginning in pre-k.

Math skills were assessed with two WJ-III/Bat-III subtests. The Applied Problems subtest measured children's ability to analyze and solve math problems using various operations (e.g., simple comparisons, counting, addition, subtraction). The Calculation subtest measured written skills using number operations (e.g., writing numbers, addition, subtraction, multiplication). Applied Problems was gathered beginning in pre-k, and Calculation was added in first grade.

Executive function was assessed using two measures, *Forward Digit Span* and *Backward Digit Span*.[viii] The Digit Span measures assess different components of children's working memory. *Forward Digit Span* tests the phonological loop component of working memory and *Backward Digit Span* tests the central executive function component of working memory. Both of these measures were added in first grade.

Behavior skills were measured based on teacher ratings using two subscales of the *Social Skills Improvement System*[ix] (SSiS). The Social Skills subscale rates behaviors that promote positive interactions while discouraging negative interactions. The Problem Behaviors subscale rates behaviors that interfere with social behavior performance or acquisition. These measures were gathered beginning in pre-k.

In addition, the *preLAS 2000*[x] was used to measure oral language proficiency in English for all children as well as in Spanish for the DLL subsample. Scores on this measure from the fall of pre-k were used as covariates in the analyses in order to examine whether differences in children's growth on the various outcome measures were related to their initial level of language proficiency (1=Non-English/Spanish speaker, 2–3=Limited English/Spanish speaker, 4–5=Fluent English/Spanish speaker). (See Table 5.)

Classroom Observations

Observations of classroom practices were conducted in classrooms attended by children in the sample each year (see Table 1). In year 1, observations were conducted in all 199 randomly-selected pre-k classrooms attended by the children in the sample. A random sample of classrooms was selected for observation in year 2 (kindergarten). In subsequent years, the classroom observation sample consisted of the classrooms attended by the study children who were in the observed kindergarten classrooms, in order to obtain a longitudinal sample associated with the classroom observations. In year 2, observations were conducted in 296 kindergarten classrooms of the 822 attended by children in the sample (representing 434 children in the sample). In year 3, observations were conducted in 296 first-grade classrooms of the 777 attended by children in the sample (representing 447 children in the sample). In year 4, observations were conducted in 280 second-grade classrooms of the 786 attended by children in the sample (representing 403 children in the sample).

Classroom observations were conducted during the second half of the school year each year. Observations typically lasted about 3 hours. Data collectors completed standard training procedures offered by the developers, along with additional field practice, and had to meet established reliability criteria prior to gathering data (i.e., 85% agreement within one point). Inter-rater reliability data were collected for approximately 20% of the observations (PK n=43, K n=59, 1st n=59, 2nd n=59), and intra-class correlations indicated that reliability was acceptable each year (Total Score=.71-.84, Emotional Support=.80-.91, Classroom Organization=.80-.85,

Instructional Support=.47-.80). (Intra-class correlations of .40–.59 are considered in the fair range, .60–.74 good, and .75–1.0 excellent. [xi])

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 (see Table 4 for an overview, including subscales and scoring). *The Classroom Assessment Scoring System*[xii] (CLASS) was used in pre-k and the CLASS K-3[xiii] in kindergarten through second grade. The CLASS/CLASS K-3 measure teachers' interactions with children in the areas of social and emotional functioning, classroom organization and management, and curriculum implementation to support cognitive and language development. The CLASS/CLASS K-3 include 10 dimensions organized into three domains. Recent research by the developers has suggested that a single total score, based on the mean of the three domain scores, can be used.[xiv] For the current study, the single total score was used for predictive analyses, with information at the domain and dimension levels included in descriptive analyses. The scale has demonstrated good interrater reliability based on reported data from the developers (mean agreement within one point=87.1%, range=78.8%–96.9%).[xii, xiii]

Parent and Teacher Surveys

Parents completed demographic surveys each year about their family and household and teachers completed online surveys that included information about their classrooms and demographic information. The parent surveys included information about parent education that was used in the current study, coded as a three-level variable (1=less than high school, 2=high school to less than Bachelor's degree, 3=Bachelor's degree or above). Parent surveys were distributed to families through the classrooms and returned in sealed envelopes to teachers for retrieval by the research team. Parent surveys were received from 91% (1,067/1,169) of participating families in year 1, 86% (888/1,034) of participating families in year 2, 85% (821/969) of participating families in year 3, and 87% (832/951) of participating families in year 4. Teachers completed online surveys about characteristics of the classroom and their background, including classroom composition (number of boys and girls in class), length of teaching experience, and degrees earned. Teachers were asked to complete the online surveys via email requests with follow up as needed, with a completion rate of 95% (189/199) in year 1, 95% (777/822) in year 2, 91% (707/777) in year 3, and 92% (727/786) in year 4.

Analysis Approach

Sample Description and Comparison

Descriptive analyses were conducted on classroom characteristics (class size, percentage of boys/girls, percentage of children with home language other than English), teacher characteristics (teaching experience, teacher gender, race/ethnicity, education level, and Georgia PSC certification in Pre-K), children's outcomes, and classroom quality for each study year. For children's outcomes, three sets of descriptive analyses were conducted: for the full sample of

children, for the DLL subsample of children assessed in English, and for the DLL subsample assessed in Spanish. Any standard scores more than 4 SD below the mean were excluded from the analyses, given that these were likely invalid scores based on the structure of the standardized, norm-referenced assessments. This situation occurred only for Spanish assessments and resulted in 0-4 scores excluded per measure, except in the case of vocabulary scores which resulted in a larger number of exclusions (14-32). Additionally, the amount of missing data and zero-order correlations among the study variables were examined.

To test for attrition bias, t-test comparisons were conducted between those remaining in the sample and not in the sample during year 4 for child and family variables included in the HLM analyses (child age, gender, race/ethnicity, English language proficiency status, IEP status, family income status).

Child Outcomes

Children's scores on each outcome from fall of pre-k through spring of second grade were analyzed to describe the development over time of children who attended Georgia's Pre-k and to explore whether child and family characteristics, pre-k program characteristics, and pre-k and subsequent classroom quality predicted the patterns of change over time. A series of hierarchical linear models (HLM) were estimated, with separate models conducted for each outcome measure. Given the greater number of repeated child outcome measures compared to previous years, more complex quadratic models were fit for all outcomes in order to allow for varying rates of change over time.

Theses models tested longitudinal growth over the four-year period for outcomes that were assessed from pre-k through second grade (WJ-III Picture Vocabulary, WJ-III Sound Awareness, WJ-III Letter-Word Identification, WJ-III Word Attack, WJ-III Applied Problems, SSiS Social Skills, and SSiS Problem Behaviors standard scores), and over the two-year period for outcomes that were added to the assessment battery in first grade (WJ-III Passage Comprehension and WJ-III Calculation standard scores, Forward Digit Span and Backward Digit Span raw scores). The Time variable represented linear effects of growth and the Time-squared variable represented quadratic effects of change in growth rates over time within these models.

The base model included three sets of covariates: 1) Child/family characteristics – gender, race/ethnicity (Hispanic/Latino, White non-Hispanic/Latino, Black non-Hispanic/Latino, Multi-racial/Other non-Hispanic/Latino), IEP status, English/Spanish *preLAS* language proficiency level at fall of pre-k, and family income (Category One vs Two); 2) Pre-k program/classroom characteristics – provider type (private setting vs public school system) and class size; and 3) Classroom quality – Pre-k classroom quality (pre-K CLASS total) and subsequent kindergarten through second-grade classroom quality (low, medium, high K-3 CLASS categorizations for each grade). Separate sets of analyses were conducted for the full sample and for the DLL subsample for outcomes measured in English and in Spanish. In addition,

child/family characteristics, pre-k program/classroom characteristics, and classroom quality were examined to determine whether they were significant predictors of differences in growth rates (linear and quadratic for outcomes assessed from pre-k through second grade and linear for first- through second-grade outcomes only).

Analyses involved grand mean centering, mean-centered covariates, accounted for missing data on covariates, and adjusted for multiple comparisons using Benjamini-Hochberg adjustments to correct for the potential false discovery rate. Post-hoc analyses were conducted for the remaining significant interactions after the adjustments, and those estimates were used in the interpretation of results and for graphing the significant interactions. Tables do not include p-values for the intercept and covariates because they were not included in the Benjamini-Hochberg adjustments. (See Appendix for further details about the analysis approach and the analysis models.)

Quality of Classroom Practices

Analyses were conducted to examine the quality of instructional practices each year from pre-k through second grade for the observed samples of classrooms attended by children in the study. Descriptive analyses were conducted for pre-k CLASS and kindergarten, first-grade, and second-grade CLASS K-3 scores, including means and frequency distributions on the total, domain, and dimension scores. A series of successive pairwise comparisons were conducted using t-tests to compare CLASS total and domain scores between grade levels from pre-k through second grade.

Results

Children's Growth over Time

Full Sample

Results from analyses of the full sample examined longitudinal growth for children who attended Georgia's Pre-K. Specifically, results for the level of performance, gains over time, and factors predicting growth rates from pre-k through second grade on language, literacy, math, behavior, and executive function skills are presented.

Descriptive Results

Results from descriptive analyses examined the level of children's performance on these measures from pre-k through second grade, based on average standard scores, in relation to population norms (see Table 6). For norm-referenced measures, the population mean of 100 represents average performance for a typical child at a given age.

For most norm-referenced measures examined longitudinally from pre-k through second grade, the average scores were at or slightly below the norming sample mean of 100 (within 0.3 SD) at the beginning of pre-k and slightly to somewhat above the mean (up to 1 SD higher) at subsequent time points. This suggests that on average, children were performing at the expected developmental level at entry into Georgia's Pre-K on most language, literacy, math, and behavior skills – vocabulary (WJ-III Picture Vocabulary), letter and word recognition (WJ-III Letter Word Identification), math problem-solving (WJ-III Applied Problems), and problem behaviors in the classroom (SSiS Problem Behaviors) – and close to the expected level on phonological awareness (WJ-III Sound Awareness) and social skills (SSiS Social Skills). For one measure of decoding and phonemic awareness skills (WJ-III Word Attack), a sizeable number of children could not perform this task prior to kindergarten.

For literacy and math measures gathered beginning in first grade (because of the developmental level required), average scores also were close to or slightly above the mean. In general, scores on most of these measures showed a pattern of decreases over the first- and second-grade periods. Executive function skills (Forward Digit Span, Backward Digit Span) showed a different pattern of consistent increases over time, although these measures were not norm-referenced.

HLM Results

The next set of results examined the patterns of longitudinal growth through second grade on language, literacy, math, executive function, and social skills for children who attended Georgia's Pre-K, after accounting for various child, family, pre-k program and classroom characteristics, and pre-k and subsequent classroom quality. [These results are based on hierarchical linear models (HLM) analyses, with the rates of change over time indicated by the coefficients for the Time and Time-squared variables in these models.] Significant increases over time on these norm-referenced measures suggest that children's growth in skills was occurring at a greater than expected rate for their age, no changes in scores suggest that growth was

occurring at the expected rate, and decreases in scores suggest that growth was occurring at a slower than expected rate based on the norming sample. However, all children were learning, so decreasing scores do not necessarily suggest children were losing skills.

For the full sample of children, significant linear and quadratic patterns of longitudinal growth were found for most language, literacy, math, and social skills measured from pre-k through second grade (see Table 7, Table 8, Figure 1, Figure 2, and Figure 3). This included phonological awareness skills (WJ-III Sound Awareness), basic reading and decoding skills (WJ-III Letter-Word Identification, WJ-III Word Attack), math problem-solving skills (WJ-III Applied Problems), and behavior skills (SSiS Social Skills). 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 decreases through second grade (i.e., smaller gains than the norming sample). For one measure of decoding skills (WJ-III Word Attack), scores showed decreases after pre-k (and given that a number of children could not perform this task prior to kindergarten, the pattern was more generally representative of an overall decrease). For SSiS problem behavior ratings, only the quadratic effect was significant; there were no overall linear changes longitudinally, but there were slightly greater decreases in scores (improvements in behavior) in pre-k and slightly greater increases in scores (declines in behavior) at the end of second grade than at other time points. For one measure of language skills (WJ-III Picture Vocabulary), there was a significant linear trend showing slight decreases in scores (lower gains than expected) throughout this period.

For most of the measures assessed beginning in first grade, significant growth effects (larger than expected gains relative to the norming sample) were found through the end of second grade (see Table 7, Table 8, Figure 1, Figure 2, and Figure 4). For the two norm-referenced measures of literacy and math skills (WJ-III Passage Comprehension, WJ-III Calculation), children showed a pattern of negative gains (lower than expected), with relatively steeper declines in scores in second grade than in first grade. For one measure of executive function (Forward Digit Span), children showed a pattern of positive growth, with relatively greater gains in first grade than in second grade. For the other measure of executive function (Backward Digit Span), there was a significant linear trend indicating increases in scores over time (although these were not norm-referenced measures).

Estimates of the annual gains on these measures based on the HLMs provide a picture of the average amount of growth by grade from fall to spring (see Table 9). For most measures assessed over this four-year time span (WJ-III Letter-Word Identification, WJ-III Applied Problems, SSiS Social Skills, SSiS Problem Behaviors), children evidenced larger gains during pre-k and somewhat lesser gains during kindergarten, with smaller than expected gains during first grade and even greater decreases during second grade. For one measure of phonological awareness skills (WJ-III Sound Awareness), children evidenced gains each year from pre-k to first grade,

although decreasing in magnitude, with declines in scores evidenced in second grade. In contrast, for one measure of decoding skills (WJ-III Word Attack), scores showed increasing declines each year. (Note that annual gains could not be calculated for WJ-III Picture Vocabulary due to model distributions.) For literacy and math measures gathered in first and second grades only (WJ-III Passage Comprehension, WJ-III Calculation), the pattern showed similarly smaller decreases in scores in first grade and greater decreases during second grade. A parallel pattern was found for the executive function measures (Forward Digit Span, Backward Digit Span), with larger gains in scores during first grade and smaller gains during second grade.

DLL Subsample

For the subsample of Spanish-speaking DLLs who attended Georgia's Pre-K Program, a similar set of results examined their longitudinal growth in both English and Spanish. Specifically, results for the level of performance, gains over time, and factors predicting growth rates from pre-k through second grade on language, literacy, math, behavior, and executive function skills are presented.

Descriptive Results - English

For children in the DLL subsample, results from descriptive analyses examined children's level of performance on these measures from pre-k through second grade. They showed somewhat similar patterns to the full sample for measures of skills in English, although their scores generally were lower (see Table 10).

For most norm-referenced language, literacy, and math measures in English, children's average scores were below the norming sample mean at the beginning of pre-k, although they increased to the mean or above in subsequent years. This suggests that on average, children in the DLL subsample were performing below the expected developmental level relative to the norming sample at entry into Georgia's Pre-K on these skills in English, but were performing around the expected level by the end of second grade. Scores were lower at pre-k entry (1.2 SD vs 0.6 SD) for phonological awareness (WJ-III Sound Awareness) than for letter and word recognition (WJ-III Letter Word Identification) and math problem-solving skills (WJ-III Applied Problems). Scores were close to or above the mean in subsequent years (0.2-0.7 SD), with slightly higher scores for letter and word recognition skills. Scores remained above the mean for decoding and phonemic awareness skills (WJ-III Word Attack), although it should be noted that many children could not perform this task prior to kindergarten. Scores on these measures generally showed patterns of increases from pre-k through kindergarten but then decreases through second grade. In contrast to other measures, scores on vocabulary skills (WJ-III Picture Vocabulary) were substantially below the mean (1.0-1.5 SD below) throughout this period, although there were slight increases over time.

For children in the DLL subsample, the pattern of scores for English skills assessed beginning in first grade also was similar to the full sample, although scores were lower overall. For the math

measure of written computation skills (WJ-III Calculation), average scores initially were somewhat above the norming sample mean (about 0.4 SD higher), with consistent decreases to slightly below the mean by the end of second grade. For written comprehension skills (WJ-III Passage Comprehension), scores were slightly below the norming sample mean at the beginning of first grade and somewhat lower through the end of second grade (0.6 SD below the mean). Executive function skills (Forward Digit Span, Backward Digit Span) showed a different pattern of consistent increases over time, with scores similar to those of the full sample, although these measures were not norm-referenced.

Descriptive Results - Spanish

For these same skills measured in Spanish, the level of performance was somewhat different (see Table 11). For most measures assessed from pre-k through second grade, the average scores at pre-k entry were well below the norming sample mean (1-2 SD) and remained below the mean in most cases. Scores tended to be lower for language and literacy skills than for math skills, with especially low scores for vocabulary (Bat-III Picture Vocabulary). Both vocabulary and letter and word recognition skills in Spanish (Bat-III Letter-Word Identification) showed a pattern of decreasing scores over time. For other measures in Spanish, phonological awareness skills (Bat-III Sound Awareness) and math problem-solving skills (Bat-III Applied Problems), scores showed patterns of gains through the earlier grades followed by decreases in first and second grades. For one measure of phonemic awareness and decoding skills (Bat-III Word Attack), average scores were somewhat higher in the earlier years (although it should be noted that many children could not perform this task prior to kindergarten).

For Spanish norm-referenced assessments beginning in first grade, the patterns were somewhat different for literacy and math skills. Scores for written comprehension skills (Bat-III Passage Comprehension) were substantially below the Spanish norming sample mean (1.3-1.9 SD), with decreases over time. In contrast, scores for written computation skills (Bat-III Calculation) remained close to the norming sample mean, with slight declines from first to second grade. Scores on the executive function measures in Spanish (Forward Digit Span, Backward Digit Span) were slightly lower than those in English, with slight increases over time (although these were not norm-referenced).

HLM Results - English

For the subsample of Spanish-speaking DLLs who attended Georgia's Pre-K Program, a similar set of results to the full sample examined the pattern of longitudinal gains in language, literacy, math, and executive function skills measured in English, after accounting for various child, family, pre-k program and classroom characteristics, and pre-k and subsequent classroom quality. (These results are based on HLM analyses, with the rates of change over time indicated by the coefficients for the Time and Time-squared variables in these models.)

Significant linear and quadratic effects were found for language and literacy skills (WJ-III Picture Vocabulary, WJ-III Sound Awareness, WJ-III Letter-Word Identification, WJ-III Word Attack) and math skills (WJ-III Applied Problems) measured in English from pre-k through second grade for the DLL subsample (see Table 12, Table 13, Figure 5, and Figure 6). Similarly to the full sample, children generally showed a pattern of gains scores during pre-k and kindergarten (greater than expected gains relative to the norming sample), with scores starting to level off in kindergarten or first grade (gains at the expected rate), and then decreases through second grade (smaller than expected gains).

For norm-referenced measures assessed in English beginning in first grade, the results were somewhat different (see Table 12, Table 13, Figure 5, Figure 6, and Figure 7). There were similar quadratic growth effects for one measure of math skills (WJ-III Calculation), indicating lower than expected gains for written computation skills, with smaller decreases in first grade and greater decreases in second grade. For the measures of literacy skills (WJ-III Passage Comprehension) and executive function (Forward Digit Span, Backward Digit Span), there were no significant growth effects, indicating that scores were constant over time.

Estimates of the annual gains on these measures based on the HLMs provide a picture of the average amount of growth by grade from fall to spring for the DLL subsample (see Table 14). Overall, children showed greater gains in pre-k and somewhat lesser gains in kindergarten. For some measures, children continued to show slight gains in first grade and then decreases in scores during second grade (WJ-III Picture Vocabulary, WJ-III Sound Awareness), whereas for others (WJ-III Letter-Word Identification, WJ-III Applied Problems) they showed increasing declines during first and second grades. For one measure (WJ-III Word Attack), gains were evidenced only in pre-k, with increasing rates of decline in subsequent grades. For literacy and math measures assessed only in first and second grades (WJ-III Passage Comprehension, WJ-III Calculation), scores showed a pattern of increasing declines from the first grade to second grade years. Similarly to the full sample, for assessments in first and second grades only, scores on language and literacy measures (WJ-III Passage Comprehension, WJ-III Calculation) showed a pattern of smaller decreases in first grade and greater decreases during second grade. For the executive function measures (Forward Digit Span, Backward Digit Span), children evidenced smaller gains in scores during first grade and larger gains during second grade.

HLM Results - Spanish

A parallel set of analyses to those conducted for the English measures examined the patterns of longitudinal growth on language, literacy, math, and executive function skills measured in Spanish for children in the DLL subsample, after accounting for various child, family, pre-k program and classroom characteristics, and pre-k and subsequent classroom quality. (These results are based on HLM analyses, with the rates of change over time indicated by the coefficients for the Time and Time-squared variables in these models.)

For norm-referenced measures in Spanish from pre-k through second grade, children in the DLL subsample displayed a mixed pattern of longitudinal growth (see Table 15, Table 16, Figure 8, and Figure 9). There were significant linear and quadratic effects of growth for all measures, although these growth patterns varied. For some measures of language and literacy skills (Bat-III Picture Vocabulary, Bat-III Letter-Word Identification, Bat-III Word Attack), scores decreased from pre-k through second grade, but the amount of decrease lessened over time. For other measures of phonological awareness and math skills (Bat-III Sound Awareness, Bat-III Applied Problems), scores increased from pre-k through kindergarten, followed by decreases from first through second grade.

For skills measured in Spanish beginning in first grade, there were no significant growth effects for children in the DLL subsample (see Table 15, Table 16, Figure 8, Figure 9, and Figure 10). For the norm-referenced measures of literacy and math skills (Bat-III Passage Comprehension, Bat-III Calculation), these results indicate that growth was at the expected level. For the measures of executive function in Spanish (Forward Digit Span, Backward Digit Span), these results indicate that scores remained constant over time.

Estimates of the annual gains on these measures based on the HLMs provide a picture of the average amount of growth in Spanish by grade from fall to spring for the DLL subsample (see Table 17). For Spanish measures of language and literacy skills (Bat-III Picture Vocabulary, Bat -III Letter-Word Identification, Bat -III Word Attack) assessed from pre-k through second grade, children showed the greatest gains in pre-k, with subsequent declines in scores but at a decreasing rate (with even slight gains by second grade in some cases). In contrast, for math skills (Bat -III Applied Problems), gains were greatest in pre-k but still positive in kindergarten, with increasing declines in scores at a similar magnitude to the gains through first and second grade. For measures assessed in first and second grades, the patterns were slightly different. For literacy skills (Bat -III Passage Comprehension), children exhibited greater declines in written comprehension scores from first grade to second grade, while for math skills (Bat -III Calculation), children showed no changes during first grade and declines during second grade in written computation skills. For the executive function measures (Forward Digit Span, Backward Digit Span), children evidenced slightly larger gains in scores during first grade and smaller gains during second grade, similar to previous patterns on these measures.

Predictors of Children's Growth – Full Sample

Child/Family Characteristics (Full Sample)

The next set of results examined specific child and family characteristics as potential predictors of children's rates of longitudinal growth for the full sample. (These results are based on HLM analyses, with the interactions with Time and Time-squared indicating predictors of rates of change in growth for that outcome.) The potential predictors included child gender, race/ethnicity (Hispanic/Latino, White non-Hispanic/Latino, Black non-Hispanic/Latino, Multi-

racial/Other non-Hispanic/Latino), IEP status, English language proficiency level at pre-k entry, and family income (Category One vs Category Two). Significant effects were found for only two of these characteristics – language proficiency and race/ethnicity. In general, the pattern of significant effects showed that children who entered pre-k with lower skill levels exhibited relatively greater rates of early growth, but often relatively lower rates in later grades than children who entered pre-k with higher skill levels.

Effects of English Proficiency Level

Children's language proficiency levels were individually assessed at the beginning of the pre-k year and categorized according to fluency levels (see Table 5). Significant effects of English language proficiency were found across several norm-referenced measures of language, literacy, math, and social skills. In most cases, children with lower levels of English proficiency at pre-k entry exhibited greater changes in growth rates, with greater initial gains but also relatively greater declines in later grades compared to children with higher levels of language proficiency. (See Table 7 and Table 8.)

There were significant effects of English language proficiency level on children's growth for three of the four measures of language and literacy skills assessed from pre-k through second grade. On WJ-III Picture Vocabulary, scores for children at the lowest proficiency level showed slightly positive growth over time through first grade followed by slight decreases, while scores for children at higher proficiency levels showed a pattern of slight decreases over time. Further, the rate of decline in vocabulary scores was relatively greater for children with lower proficiency levels (see Figure 11). On WJ-III Sound Awareness, children at all proficiency levels exhibited a general pattern of greater than expected growth (increases in standard scores) through first grade followed by lower than expected growth (declines in standard scores) in second grade. Children at the lowest proficiency level exhibited relatively greater initial increases and lesser declines later on phonological awareness skills than children at the highest proficiency level (see Figure 12). On WJ-III Letter-Word Identification, children at lower proficiency levels exhibited greater initial growth in the early grades but greater declines in later grades than children at higher levels (see Figure 13). In contrast, on WJ-III Passage Comprehension, children at all levels showed a pattern of lower than expected growth (declines in standard scores) over the first- to second-grade assessment period. Children at the next-to-lowest levels of proficiency exhibited greater decreases in these written comprehension skills than children at higher levels. (See Figure 14.)

There also were significant effects of English language proficiency for the measure of math skills assessed from pre-k through second grade. On WJ-III Applied Problems, children at all proficiency levels displayed a pattern of initial increases in scores into first grade followed by decreases in scores. Children at lower proficiency levels exhibited relatively greater initial growth on math problem-solving skills than those at higher levels, with relatively greater declines in scores in later grades for children at the lowest level compared to children at the

highest level (see Figure 15). Significant effects of English language proficiency also were found for changes in teacher ratings of social skills from pre-k through second grade. On the SSiS Social Skills measure, teachers rated children at lower proficiency levels as showing greater increases in scores in early grades and greater declines in later grades than their peers at higher proficiency levels, with the largest differences for children at the lowest proficiency level and the least differences for children at the highest proficiency level (see Figure 16).

Effects of Race/Ethnicity

Significant effects of race/ethnicity were found for most language, literacy, and math skills. In general, patterns of growth differed more for Hispanic/Latino children compared to children from other racial/ethnic groups. However, the effects of race/ethnicity and language proficiency are interrelated since almost all of the children with lower levels of English language proficiency were Hispanic/Latino. There were some similarities in the pattern of effects; in particular, a predominant effect of race/ethnicity showed that Hispanic/Latino children often exhibited greater initial gains but also greater declines in later grades compared to other groups in many cases. (See Table 7 and Table 8.)

Significant effects of race/ethnicity were found for three of the four measures of language and literacy skills assessed from pre-k through second grade. On WJ-III Picture Vocabulary, Hispanic/Latino children exhibited a different pattern of growth, with greater initial gains than expected in scores followed by smaller than expected gains in later grades. In contrast, children from other racial/ethnic groups showed a general pattern of smaller than expected gains (or declines in standard scores) in vocabulary over time, with the greatest differences in growth patterns between Hispanic/Latino and White children (see Figure 17). On WJ-III Letter-Word Identification, Hispanic/Latino children showed relatively greater initial gains but greater declines in scores over time compared to White and Black children (see Figure 18). On WJ-III Word Attack, Hispanic/Latino and White children showed overall greater growth in phonemic awareness and decoding skills than Black children; Hispanic/Latino and White children displayed slightly greater initial gains into kindergarten followed by slight declines compared to the norming sample, while Black children displayed an overall pattern of declines in scores (see Figure 19). Significant effects of race/ethnicity also were found for literacy skills assessed over the first- to second-grade period on WJ-III Passage Comprehension. All children showed slower than expected rates of acquiring written comprehension skills (declining standard scores); however, these rates were relatively slower for Black children than for Hispanic/Latino and White children (see Figure 20).

For math skills, there were significant effects from pre-k through second grade on WJ-III Applied Problems. Children from all groups displayed an initial pattern of greater than expected growth (gains in standard scores) in math problem-solving skills through kindergarten or into first grade followed by lower than expected growth (declines in standard scores) in subsequent grades. Both Hispanic/Latino and Black children showed relatively greater initial gains in scores but also

relatively greater declines in later grades than White children (see Figure 21). On WJ-III Calculation, which was assessed beginning in first grade, all groups showed generally lower than expected gains (decreasing standard scores) in written mathematical computation skills over time. Hispanic/Latino children exhibited relatively greater declines in scores than both White children and children who were Multi-racial or from other racial/ethnic groups (see Figure 22).

Pre-k Program/Classroom Characteristics (Full Sample)

Pre-K classroom and program characteristics also were examined as potential predictors of children's rates of growth from entry into Georgia's Pre-K through second grade for the full sample, after accounting for other child, family, pre-k program and classroom characteristics, and pre-k and subsequent classroom quality. (These results are based on HLM analyses, with the interactions with Time and Time-squared indicating predictors of rates of change in growth for that outcome.) These included pre-k provider type (public school system vs private program) and class size in pre-k. There were several significant effects for pre-k provider type on norm-referenced measures of children's language, literacy, math, and behavior skills. In general, children who attended Georgia's Pre-K in public school settings displayed relatively faster rates of skill acquisition in earlier grades than those in private settings, but often relatively slower rates in later grades. There were no effects for class size. (See Table 7 and Table 8.)

There were significant effects of provider type for some measures of language and literacy skills. On WJ-III Sound Awareness, children who attended public school Georgia's Pre-K settings displayed faster than expected growth (slightly greater increases in scores) from pre-k through first grade but also slower than expected growth (slightly greater declines in scores) in second grade compared to children who attended private settings (see Figure 23). Similarly, for WJ-III Letter-Word Identification, children who attended public school Georgia's Pre-K settings displayed slightly greater initial growth from pre-k through entry into first grade but also slightly slower later growth through second grade compared to children who attended private settings (see Figure 24). In contrast, on WJ-III Word Attack, children who attended public school Georgia's Pre-K settings showed slightly greater than expected growth (greater increases in standard scores) through kindergarten entry followed by slower than expected rates of growth (greater decreases) through second grade, whereas children who attended private settings showed an overall pattern of lower than expected growth (decreases in standard scores); further the rate of decline in later grades was relatively greater for those who attended public school than private settings (see Figure 25).

There was also a significant effect of provider type for one measure of math skills assessed during first and second grades. On WJ-III Calculation, children exhibited declines in written computation scores during this period, with relatively greater declines for children who attended private settings than those who attended public school settings (see Figure 26). For SSiS Problem Behaviors, children who attended public schools settings during Georgia's Pre-K were

rated by teachers as showing slightly lower scores over time compared to children who attended private settings. Further, the growth patterns differed, with children from public school settings showing slight decreases in problem behavior scores through kindergarten followed by slight increases, whereas children from private settings showed very slight increases in scores over time (see Figure 27).

Classroom Quality Predictors (Full Sample)

The next set of results examined the extent to which the quality of pre-k or later elementary school classroom instructional practices each related to children's growth in skills for the full sample. (These results are based on HLM analyses, with predictor effects indicated by the interactions with either the Time or the Time-squared variables in these models.) These analyses included CLASS scores (pre-k total score or K-2nd low, medium, high categorical scores) as the measure of classroom quality at each grade level. (See Table 7 and Table 8).

There was one significant effect for later instructional quality in elementary school (adjusting for earlier pre-k quality as well as all other child/family and pre-k program/classroom characteristics). There was a significant effect of medium classroom quality in subsequent grades on WJ-III Sound Awareness. These results indicated that children who subsequently attended elementary school classrooms rated medium quality (compared to classrooms not rated medium quality) had slightly greater growth in phonological awareness skills, although these differences were no longer apparent by the end of second grade (see Figure 28).

Predictors of Children's Growth in English – DLL Subsample

Child/Family Characteristics (DLL Subsample – English)

The next set of results examined specific child and family characteristics as potential predictors of children's rates of longitudinal growth on English measures for the DLL subsample. (These results are based on HLM analyses, with the interactions with Time and Time-squared indicating predictors of rates of change in growth for that outcome.) The potential predictors that could be examined within this subsample included child gender, English language proficiency level at pre-k entry, and family income (Category One vs Category Two). In contrast to the full sample, race/ethnicity was excluded because nearly all children were Hispanic/Latino and IEP status was excluded because very few children had IEPs. Significant effects were found for English language proficiency.

Effects of English Proficiency Level

Children's language proficiency levels were individually assessed at the beginning of the pre-k year, and categorized according to fluency levels (see Table 5). Significant effects of English language proficiency were found for the DLL subsample across some language and literacy and math skills measured in English from pre-k through second grade. (See Table 12 and Table 13.)

On WJ-III Picture Vocabulary, the pattern of gains was different for children at the lowest proficiency level compared to others. Children at the lowest proficiency level showed slightly greater than expected growth (increases in standard scores) through first grade followed by slightly less than expected growth (decreases in standard scores), while children at higher proficiency levels showed a pattern of slightly less than expected growth (decreases in scores) over time. Further, the rate of decline in vocabulary scores in later grades was relatively greater for children with lower proficiency levels (see Figure 29). On WJ-III Sound Awareness, children at all proficiency levels exhibited a general pattern of greater than expected growth in phonological awareness skills (increases in standard scores) through first grade followed by less than expected growth (decreases in standard scores) during second grade. Children at the lowest proficiency level exhibited slightly lower rates of decline in phonological awareness than children at the second highest proficiency level (see Figure 30). On WJ-III Applied Problems, children generally exhibited a pattern of greater than expected growth (increases in scores) through entry into first grade, followed by less than expected growth (decreases in scores) through the end of second grade. For children at the lowest proficiency level, the rate of change in these math problem-solving skills was significantly greater than for children at higher levels, both for initial gains and later declines (see Figure 31).

Pre-k Program/Classroom Characteristics (DLL Subsample – English)

Pre-K classroom and program characteristics also were examined as potential predictors of children's rates of growth on English outcomes from entry into Georgia's Pre-K through second grade for the DLL subsample, after accounting for other child, family, pre-k program and classroom characteristics, and pre-k and subsequent classroom quality. (These results are based on HLM analyses, with the interactions with Time and Time-squared indicating predictors of rates of change in growth for that outcome.) These included pre-k provider type (public school system vs private program) and class size in pre-k. There were no significant effects for any of these characteristics on English measures of children's language, literacy, math, or executive function skills. (See Table 12 and Table 13.)

Classroom Quality Predictors (DLL Subsample – English)

The next set of results examined the extent to which the quality of pre-k or later elementary school classroom instructional practices each related to children's growth in English skills for the DLL subsample. (These results are based on HLM analyses, with predictor effects indicated by the interactions with either the Time or the Time-squared variables in these models.) These analyses included CLASS scores (pre-k total score or K-2nd low, medium, high categorical scores) as the measure of classroom quality at each grade level. (See Table 12 and Table 13).

There was one significant effect of pre-k classroom quality on WJ-III Letter-Word Identification. Children in the DLL subsample who attended higher quality Georgia's Pre-K classrooms exhibited greater growth over time than children who attended lower quality pre-k classrooms (see Figure 32).

There also were significant effects for instructional quality in elementary school. There were two significant effects of high classroom quality in subsequent grades on language and math skills measured in English for children in the DLL subsample. On WJ-III Picture Vocabulary, children who attended elementary school classrooms rated high quality exhibited slightly greater than expected growth (increases in standard scores) from pre-k through second grade; in comparison, children who attended elementary school classrooms of lower quality (rated not high quality) began pre-k with lower scores and exhibited a slightly greater rate of growth through entry into first grade followed by slightly less than expected growth (declines in scores) through second grade (see Figure 33). On WJ-III Calculation, children in high quality classrooms exhibited slightly greater than expected growth (gains in scores) in first grade followed by slightly less than expected growth (declines in scores) in second grade, while children in not high quality classrooms exhibited less than expected growth (declines in scores) on written computation skills throughout this time period (see Figure 34).

There also were two significant effects of medium classroom quality on language and literacy skills measured in English for children in the DLL subsample. On WJ-III Picture Vocabulary, similar effects were found as for high quality. Children who attended elementary school classrooms rated medium quality exhibited slightly greater than expected growth from pre-k through second grade; in contrast, children who attended elementary school classrooms not rated medium quality began pre-k with lower vocabulary scores and exhibited a pattern of slightly greater than expected gains through first grade followed by slightly less than expected gains through second grade (see Figure 35). On WJ-III Passage Comprehension, children generally exhibited a pattern of less than expected growth (decreasing scores) over the first- and second-grade assessment period. Children who attended elementary school classrooms of medium quality exhibited a greater rate of decline in written comprehension scores than children who attended classrooms that were not medium quality (see Figure 36).

Predictors of Children's Growth in Spanish– DLL Subsample

Child/Family Characteristics (DLL Subsample – Spanish)

The next set of results examined specific child and family characteristics as potential predictors of children's rates of longitudinal growth on Spanish measures for the DLL subsample. (These results are based on HLM analyses, with the interactions with Time and Time-squared indicating predictors of rates of change in growth for that outcome.) The potential predictors that could be examined within this subsample included child gender, Spanish language proficiency level at pre-k entry, and family income (Category One vs Category Two). In contrast to the full sample, race/ethnicity was excluded because nearly all children were Hispanic/Latino and IEP status was excluded because very few children had IEPs. Significant effects were found for Spanish language proficiency.

Effects of Spanish Proficiency Level

Children's language proficiency levels were individually assessed at the beginning of the pre-k year, and categorized according to fluency levels (see Table 5). Significant effects of Spanish language proficiency were found for the DLL subsample across some measures of language and literacy skills assessed in Spanish from pre-k through second grade . (See Table 15 and Table 16.)

On Bat-III Sound Awareness, children generally showed greater than expected initial growth in Spanish phonological awareness skills followed by slight declines in later growth rates. However, this pattern was significantly different for children at the lowest proficiency level compared to children at the two highest levels. The pattern of change was more negative for children at the lowest proficiency level, with a lower rate of initial gains and an earlier and greater rate of decline in scores (see Figure 37). On Bat-III Letter-Word Identification, children at all proficiency levels exhibited a pattern of less than expected growth (decreasing scores) from pre-k through second grade; however, the rate of decline was significantly greater for children at the lowest Spanish proficiency level compared to children at the highest level (see Figure 38). Similarly, on Bat-III Word Attack, children at all Spanish proficiency levels displayed a pattern of less than expected growth (decreasing scores) over time from pre-k through second grade; however, the rate of decline was significantly greater for children at the lowest Spanish proficiency level compared to children at the two highest levels (see Figure 39).

Pre-k Program/Classroom Characteristics (DLL Subsample – Spanish)

Pre-K classroom and program characteristics also were examined as potential predictors of children's rates of growth on Spanish outcomes from entry into Georgia's Pre-K through second grade for the DLL subsample, after accounting for other child, family, pre-k program and classroom characteristics, and pre-k and subsequent classroom quality. (These results are based on HLM analyses, with the interactions with Time and Time-squared indicating predictors of rates of change in growth for that outcome.) These included pre-k provider type (public school system vs private program) and class size in pre-k. There were no significant effects for any of these characteristics on Spanish measures of children's language, literacy, math, or executive function skills. (See Table 15 and Table 16.)

Classroom Quality Predictors (DLL Subsample – Spanish)

The next set of results examined the extent to which the quality of pre-k or later elementary school classroom instructional practices each related to children's growth in Spanish skills for the DLL subsample. (These results are based on HLM analyses, with predictor effects indicated by the interactions with either the Time or the Time-squared variables in these models.) These analyses included CLASS scores (pre-k total score or K-2nd low, medium, high categorical scores) as the measure of classroom quality at each grade level. (See Table 15 and Table 16).

There were two significant effects of pre-k classroom quality on children's growth in Spanish skills for math and executive function. On Bat-III Applied Problems, children exhibited a pattern

of slightly greater than expected growth (increases in scores) in math skills measured in Spanish through first grade followed by slightly less than expected growth (decreases in scores); however, these gains and declines were relatively lower for children who attended higher quality Georgia's Pre-K classrooms compared to those who attended lower quality classrooms (see Figure 40). For one measure of executive function skills in Spanish, Forward Digit Span, children displayed an overall pattern of slightly greater than expected gains during the first- to second-grade assessment period; however, the rate of gain was slightly lower for children who had attended higher quality pre-k classrooms compared to those who had attended lower quality pre-k classrooms (see Figure 41). There were no significant effects for later classroom quality on children's growth in Spanish skills.

Quality of Classroom Practices from Pre-K through Second Grade

The quality of teacher-child instructional interactions each year from pre-k through second grade was examined using the CLASS (see Table 18). Scores were slightly higher in pre-k classrooms than in subsequent grades, although the patterns were similar. The average CLASS Total scores were in the middle range each year (4.2-4.6), although higher in earlier grades. Classroom practices were stronger in Emotional Support and Classroom Organization, with average scores in the middle to high quality range (5.1-5.7), than in Instructional Support, with average scores in the low to middle range (2.2-2.6). Average scores on the individual dimensions within each domain generally were in the same range as the overall domain scores, with a similar pattern over time.

Although the general pattern was similar across grades for the average scores, there were some differences in the distribution of individual classroom Total and domain scores. Overall, a greater proportion of pre-k classrooms scored higher on the scale (and a smaller proportion scored lower) compared to kindergarten through second-grade classrooms. On the CLASS Total score, most classrooms scored in the middle range (3.0-4.9) of the scale (PK=73%, K=79%, 1st=87%, 2nd=80%). However, a higher proportion of pre-k classrooms scored in the high range (5.0-7.0) of the scale (PK=26%, K=17%, 1st=9%, 2nd=13%); further, relatively fewer pre-k classrooms scored at the lowest end of that range. Few classrooms scored in low range (1.0-2.9) of the scale at any grade, with almost no pre-k classrooms scoring in this range on the Total. (See Figure 42.)

On Emotional Support, almost all pre-k classrooms scored in the high (5.0-7.0) range, compared to about two-thirds of the subsequent grades (PK=89%, K=68%, 1st=66%, 2nd=73%); further, substantially more pre-k classrooms scored in the highest range (6.0-7.0) of the scale (PK=44%, K=17%, 1st=9%, 2nd=18%). In contrast, fewer pre-k classrooms scored in the middle range (3.0-4.9) of the scale (PK=11%, K=31%, 1st=34%, 2nd=27%); almost no classrooms scored in the lower range (1.0-2.9) at any grade. (See Figure 43.)

A similar pattern was found for Classroom Organization, although the differences were not as large. (See Figure 44.) The majority of classrooms scored in the upper range (5.0-7.0) overall (PK=78%, K=72%, 1st=61%, 2nd=64%), with a slightly higher proportion of pre-k classrooms scoring in the highest range (6.0-7.0) of the scale (PK=25%, K=19%, 1st=10%, 2nd=19%). Similarly, slightly fewer pre-k classrooms scored in the middle range (3.0-4.9) of the scale (PK=22%, K=27%, 1st=37%, 2nd=36%), and almost no classrooms scored in the lower range (1.0-2.9) at any grade.

In contrast, scores were distributed toward the lower end of the scale for all grades on Instructional Support. (See Figure 45.) The majority of classrooms scored in the lower range (1.0-2.9) of the scale (PK=74%, K=72%, 1st=80%, 2nd=86%). However, slightly fewer pre-k classrooms scored at the lowest range (1.0-1.9) compared to older grades (PK=22%, K=29%, 1st=30%, 2nd=38%). A slightly higher proportion of pre-k and kindergarten classrooms scored in the middle range (3.0-4.9) of the scale compared to first- and second-grade classrooms (PK=26%, K=28%, 1st=20%, 2nd=14%), and very few classrooms scored in the high range (5.0-7.0) of the scale at any grade.

Comparisons of CLASS results each year for the observed sample of classrooms attended by children in the study indicated that the Georgia's Pre-K classrooms generally scored higher than subsequent grades (based on t-test comparisons). These results were found for the Total score as well as for the three domain scores in most cases (see Table 18). Scores were significantly higher in the Georgia's Pre-K classrooms than in the kindergarten classrooms for the CLASS Total and two of the three domains (Emotional Support, Classroom Organization). Georgia's-Pre-K classrooms also scored significantly higher than the first- and second-grade classrooms on the CLASS Total and all three domains (Emotional Support, Classroom Organization, Instructional Support). There were fewer differences for kindergarten classrooms relative to older grades. Kindergarten classrooms scored significantly higher than first-grade classrooms on the CLASS Total and one domain (Classroom Organization) and higher than second-grade classrooms on one domain (Instructional Support). There were significant differences between first-grade and second-grade classrooms on two domains, but in opposite directions; first-grade classrooms scored significantly higher on Instructional Support and second-grade classrooms scored significantly higher on Emotional Support.

Summary and Conclusions

This ongoing longitudinal study is following a sample of over 1,100 children who attended Georgia's Pre-K Program from pre-k through fifth grade to examine factors related to school readiness and later school success. These results focused on children's longitudinal performance on a variety of measures of language, literacy, math, executive function, and behavior skills from pre-k (year 1) through second grade (year 4), as well as the quality of their educational experiences each year. Overall, these findings suggest that children who attended Georgia's Pre-K exhibited a general pattern of initial gains for most skills from pre-k through kindergarten or into first grade, followed by increasing declines in scores through second grade. These initial gains in standard scores suggest that children were growing at greater than expected rates for typical development, while the declines in standard scores suggest that they were gaining at lower than expected rates. In other words, children were gaining more than a school year's worth of knowledge during pre-k and kindergarten, about a school year's worth in first grade, but often less than a school year's worth in second grade on these measures. Although children were continuing to learn and develop throughout this time period, the decreasing rate of growth during the later elementary school years suggests that the early benefits of pre-k for enhancing growth trajectories were not being sustained.

In general, children's scores were near or slightly below the population average when they entered Georgia's Pre-K, and scores remained in this range by the end of second grade. This is based on results from a variety of norm-referenced measures of language, literacy, math, and behavior skills assessed at the beginning and end of each school year from pre-k through second grade. However, children often exhibited non-linear (or quadratic) patterns of growth on these measures. They made greater gains earlier (during pre-k and kindergarten), with the rate of gain lessening over time, and began to show increasing declines in scores later (from first to second grade). This pattern was found for a range of language and literacy skills (phonological awareness, letter and word recognition), math skills (math problem-solving), and behavior skills (social skills). For two language and literacy skills (vocabulary and decoding/phonemic awareness), there were overall longitudinal decreases in scores.

For more advanced literacy and math skills assessed only in first and second grades (written comprehension, written computation), children showed a similar pattern of declining scores during this time period as well. In contrast, for measures of executive function assessed in first and second grades (which were based on raw scores), they showed positive growth during that time period.

These overall findings of initial effects related to pre-k participation that are not sustained over time are consistent with results from most studies of the longer-term impacts of early childhood programs. A meta-analysis of experimental studies of early childhood programs found gains by the end of the program (average of about one-third SD), but those gains generally were not

maintained into elementary school[xv]. Similarly, a recent synthesis of pre-k program evaluations showed substantial gains during the pre-k year on academic skills, with much smaller differences between those who had and had not attended the programs by third grade[xvi].

In this study, growth in skills for Spanish-speaking DLLs was measured in both English and Spanish using parallel measures. The results for language, literacy, and math skills assessed in English were similar for the DLL subsample to those for the full sample, with a general pattern of greater earlier gains during pre-k and kindergarten followed by increasing declines through second grade. These findings are consistent with other studies of many pre-k evaluations that document larger gains by DLLs, at least during the pre-k year[xvi]. In contrast, for most language and literacy skills assessed in Spanish, they exhibited decreases in scores over time, although the amount of decrease lessened over time. For a few skills (phonological awareness, math problem-solving), they showed similar patterns to the English measures. For more advanced reading and math skills assessed beginning in first grade, they showed no changes in scores, indicating that growth was at the expected rate, although their scores were low overall for most Spanish skills. These results suggest that although children in the DLL subsample were continuing to develop across all domains of learning in English, they were not keeping pace with regard to some of these same skills in Spanish. It is important to note that the language of instruction in these classrooms was primarily English, and there may be few resources and little support for children's home language within their classroom and school settings. However, the results from the current study suggest that pre-k seems especially effective for children with lower levels of skills in their second language, although those gains may not be maintained over time.

In terms of factors that predicted learning outcomes for children who attended Georgia's Pre-K, the strongest effects were found for children's language proficiency levels. Some effects also were found for children's race/ethnicity, pre-k provider type, and pre-k and subsequent classroom quality. Children with lower levels of English proficiency made greater initial gains, but often showed relatively lower rates of growth in later grades on several measures of language, literacy, math, and social skills compared to more language proficient peers. Similar effects were found for children in the DLL subsample for a few language and math skills measured in English. Relatively greater losses in some Spanish language and literacy skills over time were found for children with lower proficiency levels when Spanish language proficiency was considered. Children's race/ethnicity also was a significant predictor of growth for most language and literacy skills, with relatively greater initial gains followed by relatively lower rates of growth for Hispanic/Latino children than for children from other racial/ethnic groups. Not surprisingly, this pattern was similar to the pattern for language proficiency, given that nearly all children in the DLL subsample were Hispanic/Latino. These findings suggest that children's language proficiency, both in their home language and the language of instruction, has some similar effects on their learning trajectories, particularly in the later grades. Children who entered pre-k with lower language proficiency and/or DLLs gained many skills at a faster rate, thus catching up to their peers more quickly; however, as skills become more advanced in

the elementary grades, they gain at a slower rate than their peers and may start becoming even further behind. This pattern has important implications for the development of critical academic skills such as those related to later reading success.

This study extends the findings from pre-kindergarten evaluations that have similarly documented larger gains for DLL children[xvi]. This study measured the English proficiency of all children at pre-k entry, including those in the DLL subsample, and therefore was able to describe the extent to which language proficiency predicted short- and longer-term gains. These results showed that Georgia's Pre-K was especially successful at improving the skills of children with lower initial levels of English proficiency – something that is often assumed in the literature, but this study was able to demonstrate, at least in the short-term. Furthermore, unlike most studies, this study examined DLL children's skills in both their home language, Spanish, as well as in their second language, English.

For pre-k provider type, the patterns again seemed to mirror these effects for language proficiency. Children who attended public school Georgia's Pre-K sites showed greater initial gains for some language and literacy skills followed by relatively lower rates of gain through second grade, compared to children who attended private pre-k sites. In contrast, for a few math and behavior outcomes, there were some slight positive effects of attending public school pre-k sites. It is worth noting the similarity of many of the effects across various factors that were examined – for children with lower language proficiency, Hispanic/Latino children, and children who attended public school pre-k – particularly for language and literacy outcomes. Taken together, these results offer further evidence that some more vulnerable groups of children may be experiencing patterns of decreasing learning gains that are not being supported by optimal learning opportunities within the school setting. Prior studies have not examined whether pre-k impacts varied depending on whether they were located in public schools or private child care settings. These findings have important implications not only for the understanding of pre-k in Georgia but also for the rest of the country.

The quality of children's educational experiences, as measured by the CLASS, was slightly higher in the randomly-selected sample of Georgia's Pre-K classrooms attended by children in the study compared to their subsequent elementary school classrooms. However, this pattern was similar across grades, with scores generally in the middle to high range for the CLASS Total, Emotional Support, and Classroom Organization, and in the low to middle range for Instructional Support. It is important to acknowledge potential differences in the relevance of various aspects of this measure for educational practices at different ages. However, these results can be examined in conjunction with the results on child outcomes, which showed generally greater growth in pre-k as well as kindergarten compared to later years, with declining rates of growth in first and second grade in many instances. In addition, there were some effects of classroom quality, especially for DLLs, indicating generally positive effects of either pre-k or subsequent classroom quality for English language and literacy skills (vocabulary, letter and

word recognition) as well as math skills (calculation, math problem-solving). The effects on executive function skills were less clear-cut, with some indication that higher quality classrooms may be less beneficial for DLLs when measured in Spanish.

Taken as a whole, these findings suggest that the subsequent teaching and learning opportunities available to children may not be optimal for sustaining their early gains from pre-k, with these effects further magnified for children with lower proficiency levels and DLLs. Although it is not possible to establish a clear causal link between program participation and outcomes without a comparison group, this study was able to examine patterns of developmental growth over time and factors related to greater growth. Although children performed near or slightly above average levels for most skills, they showed increasingly greater declines in the level of growth in many areas by second grade. For Spanish-speaking DLLs, this pattern was even more noticeable with regard to the loss of skills in their home language, thus suggesting that instruction to support skill development both in English and their home language may be of value.

Overall, these findings suggest that children who attended Georgia's Pre-K Program experienced positive early gains in language, literacy, math, executive function, and social skills. These results appeared stronger during the pre-k and initial elementary school years, with subsequent declines in the level of gains through second grade for many outcomes. Therefore, to sustain the positive effects of children's early pre-k experiences, it would be important to ensure that policies and practices are aligned among early and later educational settings in ways that support high quality instruction and promote optimal development for all children.

Table 1. Study Sample by Grade and Year

	Pre-K Year 1 (2013-2014)	Kindergarten Year 2 (2014-2015)	1st Grade Year 3 (2015-2016)	2nd Grade Year 4 (2016-2017)
Classrooms (n)				
Full Sample	199	822	777	786
Observed Sample	199	296	296	280
Children (n)				
Full Sample	1,169	1,034	969	951
DLL Subsample	139	118	119	116

Table 2. Characteristics of Classrooms and Teachers (Pre-K – 2nd Grade)

Measure	Pre-K		Kindergarten		1st grade		2nd grade	
	n	Mean/% Range	n	Mean/% Range	n	Mean/% Range	n	Mean/% Range
Classroom Characteristics								
Class size	189	21.1 14-23	772	20.1 4-27	694	20.7 7-30	728	21.0 3-29
Percentage of boys	189	50.1% 18.2%-81.8%	768	51.3% 0.0%-100%	695	51.0% 0.0%-80%	725	51.0% 0.0%-100%
Percentage of children with home language other than English	199	11.2% 0.0%-100.0%	770	14.7% 0.0%-100.0%	687	12.0% 0.0%-100.0%	700	11.0% 0.00%-100%
Teacher Characteristics								
Years teaching current grade level	184	5.8 0-23	777	8.5 0-39	695	6.8 0-41	728	5.9 0-31
Years teaching at current school	184	5.1 0-21	774	8.8 0-33	693	8.5 0.5-37	729	8.5 0-49
Total years teaching	179	11.1 1-38	777	14.5 0-46	688	14.2 0.5-46	728	13.7 0.5-49
Female	186	98.4%	762	98.1%	688	99.1%	708	97.1%
Race/Ethnicity								
White	128	67.4%	596	76.7%	492	69.6%	523	71.9%
Black	52	27.4%	143	18.4%	164	23.2%	168	23.1%
Hispanic/Latino	5	2.7%	13	1.7%	16	2.3%	16	2.2%
Other	5	2.6%	25	3.2%	35	5.0%	20	2.8%
Georgia PSC Certification	162	81.8%	-	-	-	-	-	-
Highest Degree Earned								
Associate's Degree	4	2.2%	1	0.1%	2	0.3%	0	0.0%
Bachelor's Degree	118	64.1%	293	37.8%	245	35.3%	275	37.7%
Master's Degree	52	28.3%	338	43.5%	300	43.2%	446	61.2%
Education Specialist	9	4.9%	135	17.4%	138	19.9%	0	0.0%
PhD/EdD/PsyD	1	0.5%	10	1.2%	9	1.3%	8	1.1%

Table 3. Characteristics of Children in the Sample (2nd Grade)

	Original Year 1 Sample			
	In Year 4 Sample n=951		Not in Year 4 Sample n=218	
Characteristics[a]	Mean/%	n	Mean/%	n
Child Characteristics				
Child's age on 9/1 of second grade year	7.5	951	7.5	218
Gender				
Male	49.2%	468	45.9%	100
Female	50.8%	483	54.1%	118
Race/Ethnicity				
White	41.7%	397	41.5	91
Black/African American	37.0%	352	37.3%	81
Hispanic/Latino	15.0%	143	12.4%	27
Multi-racial/Other	6.3%	59	8.8%	19
Limited English language proficiency	10.0%	95	9.6%	21
Individualized Education Program (IEP)	2.8%	27	3.2%	7
Family Characteristics				
Income[b]				
Category One	53.8%	512	56.9%	124
Category Two	46.2%	439	43.1%	94
Primary caregiver education[c]				
< High School	10.2%	94	12.2%	26
High School - < BA/BS	61.9%	569	65.7%	140
BA/BS or above	27.9%	256	22.1%	47

[a] Source of data: Bright from the Start: Georgia Department of Early Care and Learning (DECAL) for all characteristics except parent education, which was obtained from parent surveys.

[b] Category One represents participation in one or more of the following programs: SNAP, TANF, SSI, CAPS, Medicaid, free or reduced-price meals.

[c] Data were not reported for 37 children.

Table 4. Child Outcome and Classroom Quality Measures

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

[a] WJ-III = *Woodcock Johnson Tests of Achievement*; Bat-III = *Batería III Woodcock-Muñoz Pruebas de Aprovechamiento*; SSiS = *Social Skills Improvement System*; CLASS = *Classroom Assessment Scoring System.*

[b] The English language versions of these measures were used with all children, and both the English and Spanish language versions of these measures were used with the DLL subsample.

[c] These measures are teacher ratings of individual children's skills.

[d] This measure is a classroom observation conducted by the researchers.

Table 5. Child Language Proficiency Levels at Pre-K Entry

	English				Spanish	
	Full Sample		DLL Subsample		DLL Subsample	
preLAS Proficiency Level	%	n	%	n	%	n
Level 1 (Non-Speakers)	7.9	91	47.1	65	34.1	46
Level 2 (Limited Speakers)	4.1	48	13.8	19	7.4	10
Level 3 (Limited Speakers)	14.3	166	18.1	25	17.0	23
Level 4 (Fluent Speakers)	26.2	303	11.6	16	21.5	29
Level 5 (Fluent Speakers)	47.5	551	9.4	13	20.0	27
Total	100.0	1159	100.0	138	100.0	135

Table 6. Child Outcomes for Full Sample (Pre-K – 2nd Grade)

	Pre-K				Kindergarten				1st Grade				2nd Grade			
	Fall		Spring		Fall		Spring		Fall		Spring		Fall		Spring	
Measure	n	Mean (SD)	n	Mean (SD)	n	Mean (SD)	n	Mean (SD)	n	Mean (SD)	n	Mean (SD)	n	Mean (SD)	n	Mean (SD)
Language & Literacy																
WJ-III Picture Vocabulary[a,b]	1,154	100.2 (12.8)	1,048	99.9 (11.7)	1,024	99.3 (10.6)	1,007	98.8 (10.2)	964	98.1 (10.4)	942	97.6 (10.5)	886	97.8 (10.5)	932	96.7 (10.0)
WJ-III Sound Awareness[a,b]	1,137	95.9 (17.4)	1,041	102.3 (18.8)	1,022	106.9 (19.0)	1,001	114.4 (19.7)	963	116.6 (18.1)	936	116.9 (18.0)	886	114.2 (17.5)	931	110.9 (16.6)
WJ-III Letter-Word Identification[a,b]	1,156	100.6 (13.8)	1,048	103.2 (12.7)	1,024	107.3 (12.7)	1,007	113.9 (13.2)	963	112.0 (12.9)	942	112.2 (11.8)	885	109.2 (11.1)	932	107.1 (10.6)
WJ-III Word Attack[i,b]	451	118.8 (10.9)	737	114.3 (10.6)	975	112.1 (12.3)	1,006	114.5 (12.1)	964	110.7 (10.2)	942	109.5 (10.2)	886	106.9 (10.5)	932	105.0 (10.3)
WJ-III Passage Comprehension[a,b]	-	-	-	-	-	-	-	-	962	102.3 (13.4)	940	101.2 (11.1)	886	98.6 (10.8)	932	96.6 (10.6)
Math																
WJ-III Calculation[a,b]	-	-	-	-	-	-	-	-	964	107.2 (15.5)	940	107.9 (13.9)	886	103.3 (12.8)	932	102.3 (13.9)
WJ-III Applied Problems[a,b]	1,150	102.8 (13.2)	1,049	103.7 (11.8)	1,025	104.7 (11.2)	1,006	106.5 (12.0)	960	105.8 (11.9)	937	104.7 (12.4)	885	103.1 (12.4)	932	102.7 (12.8)
Executive Function																
Forward Digit Span[c]	-	-	-	-	-	-	-	-	963	4.3 (0.8)	942	4.5 (0.8)	886	4.6 (0.9)	932	4.7 (0.8)
Backward Digit Span[c]	-	-	-	-	-	-	-	-	962	2.2 (0.6)	942	2.3 (0.7)	886	2.5 (0.6)	932	2.6 (0.7)
Behavior Skills																
SSiS Social Skills[a]	1,089	96.4 (15.9)	947	100.0 (14.9)	865	100.7 (14.5)	851	102.4 (15.1)	809	100.6 (14.8)	870	100.6 (15.3)	833	100.3 (15.5)	869	99.9 (15.8)
SSiS Problem Behaviors[a]	1,094	100.8 (15.4)	951	100.8 (15.4)	873	99.0 (13.7)	853	99.5 (14.4)	809	99.3 (14.4)	876	100.2 (14.5)	840	99.3 (14.3)	870	101.0 (14.9)

[a] Indicates standard score on norm-referenced measure with mean=100, SD=15.

[b] Scores reflect use of updated normative tables (2007).

[c] Possible range=1-8.

Table 7. Full Sample HLM Results—Language and Literacy Skills (Pre-K – 2nd Grade)

	WJ-III Picture Vocabulary		WJ-III Sound Awareness		WJ-III Letter-Word Identification		WJ-III Word Attack		WJ-III Passage Comprehension	
Effect	Est[a]	(SE)	Est[a]	(SE)	Est[a]	(SE)	Est[a]	(SE)	Est[a]	(SE)
Intercept	100.03	(0.25)	94.88	(0.51)	99.48	(0.40)	113.15	(0.39)	102.80	(0.42)
Child/Family Characteristics										
Gender[b]	1.39	(0.49)	-0.56	(0.85)	-0.83	(0.70)	-0.95	(0.74)	-1.16	(0.68)
Race										
Black	-1.70	(0.58)	-2.88	(1.07)	4.39	(0.87)	2.14	(0.88)	1.31	(0.81)
Hispanic	-9.61	(0.88)	-5.34	(1.58)	-0.78	(1.29)	0.72	(1.43)	1.28	(1.23)
Multi/Other	-1.72	(1.03)	2.05	(1.81)	6.05	(1.50)	5.23	(1.46)	3.39	(1.47)
White	1.91	(1.04)	-1.92	(1.82)	-5.96	(1.50)	-5.20	(1.46)	-5.20	(1.46)
English Proficiency										
Level 1	-26.26	(1.12)	-26.42	(1.98)	-13.34	(1.61)	-5.35	(2.03)	-10.37	(1.60)
Level 2 or 3	-8.77	(0.71)	-17.64	(1.24)	-8.25	(1.02)	-4.97	(1.13)	-5.28	(0.99)
Level 4	-5.27	(0.60)	-11.44	(1.05)	-5.37	(0.86)	-3.30	(0.92)	-4.52	(0.84)
Level 5	23.11	(1.07)	23.01	(1.88)	12.19	(1.52)	6.11	(1.86)	9.21	(1.50)
IEP[c]	-1.04	(1.50)	-6.90	(2.65)	-3.59	(2.14)	-5.28	(2.35)	-4.90	(2.07)
Family Income[d]	-2.04	(0.52)	-3.69	(0.93)	-3.95	(0.76)	-3.03	(0.79)	-2.83	(0.73)
Pre-K Characteristics										
Provider Type[e]	-0.66	(0.50)	-2.34	(1.02)	-2.08	(0.81)	-1.97	(0.75)	-1.55	(0.70)
Class Size	0.06	(0.18)	0.38	(0.36)	0.45	(0.28)	0.56	(0.28)	0.84	(0.25)
Classroom Quality										
Pre-K CLASS Total	-0.25	(0.48)	0.35	(0.99)	-0.33	(0.74)	-0.61	(0.74)	0.91	(0.55)
K-2 CLASS High	0.58	(0.59)	-1.10	(1.19)	0.43	(0.77)	1.55	(1.04)	0.33	(0.79)
K-2 CLASS Medium	0.69	(0.49)	-3.51	(1.02)	-1.40	(0.67)	-1.57	(0.92)	0.35	(0.70)
Time	-0.29*	(0.11)	9.29***	(0.21)	5.99***	(0.14)	0.55*	(0.18)	-1.42***	(0.29)
Time-squared	-0.03	(0.01)	-1.01***	(0.03)	-0.74***	(0.02)	-0.26***	(0.02)	-0.24**	(0.09)
Time x Gender[b]	0.24	(0.19)	0.35	(0.34)	-0.01	(0.22)	0.03	(0.32)	0.15	(0.17)
Time x Race	***H>M,W,B		ns		***H>W>B		**W,H>B		*** B<W,H	
Time x Black	0.02	(0.22)	0.22	(0.41)	-0.71*	(0.26)	-1.07*	(0.38)	-1.67***	(0.20)
Time x Hispanic	1.60***	(0.34)	1.41	(0.62)	1.45**	(0.40)	0.30	(0.62)	-0.74	(0.30)
Time x Multi/Other	0.14	(0.40)	1.47	(0.73)	0.02	(0.47)	-0.88	(0.64)	-0.76	(0.36)
Time x White	-0.18	(0.41)	-1.46	(0.73)	-0.03	(0.47)	0.89	(0.64)	0.76	(0.36)

[a] Significance levels are *p< .05, **p< .01, ***p< .001.
[b] Female=0, Male=1.
[c] No IEP=0, IEP=1.
[d] Category Two=0, Category One=1.
[e] Private site=0, Public school site=1.

Table 7 (Cont.) Full Sample HLM Results - Language and Literacy Skills (Pre-K – 2nd Grade)

	WJ-III Picture Vocabulary		WJ-III Sound Awareness		WJ-III Letter-Word Identification		WJ-III Word Attack		WJ-III Passage Comprehension	
Effect	Est[a]	(SE)	Est[a]	(SE)	Est[a]	(SE)	Est[a]	(SE)	Est[a]	(SE)
Time x English Proficiency	***1,2/3>4>5		ns		***1,2/3>5		ns		** 3<4,5	
Time x Level 1	4.53***	(0.44)	-1.11	(0.80)	1.78**	(0.51)	-1.34	(0.91)	-0.72	(0.40)
Time x Level 2 or 3	0.71*	(0.27)	0.25	(0.49)	0.99*	(0.32)	-0.14	(0.49)	-1.03***	(0.24)
Time x Level 4	0.74*	(0.23)	0.76	(0.42)	0.52	(0.27)	0.12	(0.40)	-0.06	(0.21)
Time x Level 5	-4.11***	(0.41)	1.42	(0.75)	-1.69***	(0.48)	0.48	(0.83)	0.76	(0.36)
Time x IEP[c]	-1.09	(0.57)	-1.29	(1.06)	-0.52	(0.65)	0.16	(1.02)	0.58	(0.50)
Time x Income	0.10	(0.20)	0.00	(0.37)	0.28	(0.23)	0.21	(0.34)	0.30	(0.18)
Time x Provider Type[e]	0.42	(0.19)	0.94*	(0.36)	0.79**	(0.24)	0.84*	(0.33)	0.01	(0.17)
Time x Class Size	0.07	(0.07)	0.14	(0.13)	0.07	(0.08)	-0.05	(0.12)	-0.13	(0.06)
Time x Pre-K CLASS	0.07	(0.20)	0.73	(0.39)	0.50	(0.25)	0.31	(0.33)	-0.28	(0.14)
Time x CLASS High	-0.61	(0.38)	1.46	(0.72)	-0.10	(0.46)	-0.20	(0.62)	0.23	(0.39)
Time x CLASS Medium	-0.38	(0.35)	2.01*	(0.66)	0.98	(0.42)	1.15	(0.58)	-0.09	(0.37)
Time-sq x Gender[b]	-0.01	(0.03)	-0.03	(0.05)	0.02	(0.03)	0.03	(0.04)	-	-
Time-sq x Race	*H<W		ns		**H<W,B		ns			
Time-sq x Black	-0.03	(0.03)	-0.07	(0.05)	-0.01	(0.03)	0.03	(0.05)	-	-
Time-sq x Hispanic	-0.13*	(0.05)	-0.13	(0.08)	-0.17**	(0.05)	-0.05	(0.08)	-	-
Time-sq x Multi/Other	0.01	(0.06)	-0.23	(0.10)	-0.04	(0.06)	0.06	(0.08)	-	-
Time-sq x White	0.00	(0.06)	0.23	(0.10)	0.05	(0.06)	-0.06	(0.08)	-	-
Time-sq x English Prof	***1<2/3,4<5		*1>5		**1<2/3,4<5		ns			
Time-sq x Level 1	-0.43***	(0.06)	0.28*	(0.11)	-0.22***	(0.07)	0.14	(0.11)	-	-
Time-sq x Level 2 or 3	-0.10	(0.04)	0.08	(0.07)	-0.09	(0.04)	0.04	(0.06)	-	-
Time-sq x Level 4	-0.10*	(0.03)	-0.04	(0.06)	-0.03	(0.03)	0.00	(0.05)	-	-
Time-sq x Level 5	0.41***	(0.06)	-0.31*	(0.10)	0.20*	(0.06)	-0.05	(0.10)	-	-
Time-sq x IEP[c]	0.17	(0.08)	0.20	(0.14)	0.07	(0.08)	-0.03	(0.13)	-	-
Time-sq x Income[d]	-0.02	(0.03)	0.02	(0.05)	0.00	(0.03)	0.02	(0.04)	-	-
Time-sq x Provider Type[e]	-0.06	(0.03)	-0.12*	(0.05)	-0.10**	(0.03)	-0.11*	(0.04)	-	-
Time-sq x Class Size	-0.02	(0.01)	-0.02	(0.02)	-0.01	(0.01)	0.00	(0.02)	-	-
Time-sq x Pre-K CLASS	0.00	(0.03)	-0.10	(0.05)	-0.06	(0.03)	-0.03	(0.04)	-	-
Time-sq x CLASS High	0.09	(0.05)	-0.19	(0.10)	0.04	(0.06)	0.01	(0.08)	-	-
Time-sq x CLASS Medium	0.04	(0.05)	-0.23	(0.09)	-0.09	(0.06)	-0.14	(0.08)	-	-

[a] Significance levels are *p< .05, **p< .01, ***p< .001.

[b] Female=0, Male=1.

[c] No IEP=0, IEP=1.

[d] Category Two=0, Category One=1.

[e] Private site=0, Public school site=1.

Table 8. Full Sample HLM Results —Math, Executive Function, and Behavior Skills (Pre-K – 2nd Grade)

	Math				Executive Function				Behavior Skills			
	WJ-III Calculation		WJ-III Applied Problems		Forward Digit Span		Backward Digit Span		SSiS Social Skills		SSiS Problem Behaviors	
Effect	Est[a]	(SE)	Est[a]	(SE)	Est[a]	(SE)	Est[a]	(SE)	Est[a]	(SE)	Est[a]	(SE)
Intercept	107.29	(0.57)	102.59	(0.33)	4.25	(0.03)	2.19	(0.03)	97.54	(0.53)	100.41	(0.55)
Child/Family Characteristics												
Gender[b]	0.16	(0.85)	0.15	(0.57)	-0.05	(0.05)	0.04	(0.04)	1.15	(0.8)	1.33	(0.78)
Race												
Black	1.01	(1.03)	-5.92	(0.70)	0.25	(0.06)	-0.03	(0.05)	-1.24	(1.03)	-0.31	(1.02)
Hispanic	5.55	(1.53)	-3.89	(1.05)	0.05	(0.09)	0.03	(0.07)	4.31	(1.51)	-5.80	(1.48)
Multi/Other	4.56	(1.82)	-0.52	(1.21)	0.23	(0.11)	0.14	(0.08)	-0.48	(1.74)	-1.27	(1.69)
White	-4.52	(1.83)	0.66	(1.21)	-0.23	(0.11)	-0.14	(0.08)	0.55	(1.74)	1.25	(1.69)
English Proficiency												
Level 1	-8.19	(2.00)	-23.68	(1.32)	-0.76	(0.12)	-0.32	(0.09)	-14.04	(1.86)	6.72	(1.81)
Level 2 or 3	-7.19	(1.23)	-12.00	(0.83)	-0.40	(0.07)	-0.26	(0.06)	-7.56	(1.18)	4.54	(1.15)
Level 4	-4.13	(1.05)	-6.91	(0.70)	-0.34	(0.06)	-0.17	(0.05)	-4.90	(0.99)	2.16	(0.97)
Level 5	7.05	(1.87)	21.22	(1.25)	0.58	(0.11)	0.29	(0.09)	12.78	(1.75)	-5.35	(1.71)
IEP[c]	-11.20	(2.55)	-4.61	(1.76)	-0.38	(0.15)	-0.21	(0.12)	-9.77	(2.49)	8.27	(2.42)
Family Income[d]	-3.20	(0.91)	-2.30	(0.62)	-0.05	(0.05)	-0.07	(0.04)	-2.07	(0.88)	1.68	(0.87)
Pre-K Characteristics												
Provider Type[e]	-1.92	(0.90)	-1.01	(0.66)	-0.03	(0.05)	-0.03	(0.04)	-1.34	(1.05)	0.16	(1.09)
Class Size	0.19	(0.32)	0.34	(0.23)	-0.01	(0.02)	0.00	(0.01)	0.08	(0.36)	-0.30	(0.38)
Classroom Quality												
Pre-K CLASS Total	0.52	(0.71)	0.71	(0.65)	0.09	(0.04)	0.02	(0.03)	0.53	(1.07)	-0.98	(1.08)
K-2 CLASS High	1.27	(1.23)	-0.65	(0.81)	-0.04	(0.07)	0.09	(0.07)	-3.42	(1.42)	2.17	(1.35)
K-2 CLASS Medium	-0.95	(1.09)	-0.55	(0.69)	0.08	(0.06)	-0.02	(0.06)	-3.90	(1.21)	1.89	(1.17)
Time	-0.75	(0.47)	1.84***	(0.15)	0.27***	(0.03)	0.16***	(0.03)	2.10***	(0.26)	-0.53	(0.24)
Time-squared	-0.42*	(0.15)	-0.29***	(0.02)	-0.05***	(0.01)	-0.01	(0.01)	-0.27***	(0.03)	0.09*	(0.03)
Time x Gender[b]	0.39	(0.27)	0.09	(0.24)	0.03	(0.02)	-0.01	(0.02)	-0.46	(0.42)	-0.74	(0.38)
Time x Race	***W,M>H		*** H,B>W		ns		ns		ns		ns	
Time x Black	-0.76	(0.32)	1.23***	(0.29)	-0.04	(0.02)	-0.03	(0.02)	-0.09	(0.50)	-0.50	(0.47)
Time x Hispanic	-1.76**	(0.49)	2.02***	(0.44)	-0.06	(0.03)	0.00	(0.03)	-0.35	(0.74)	-0.90	(0.68)
Time x Multi/Other	0.39	(0.58)	1.28	(0.52)	-0.04	(0.03)	-0.02	(0.04)	0.40	(0.91)	-0.53	(0.83)
Time x White	-0.40	(0.58)	-1.32	(0.52)	0.04	(0.03)	0.02	(0.04)	-0.43	(0.91)	0.54	(0.83)

[a] Significance levels are *p< .05, **p< .01, ***p< .001.
[b] Female=0, Male=1.
[c] No IEP=0, IEP=1.
[d] Category Two=0, Category One=1.
[e] Private site=0, Public school site=1.

Table 8 (Cont.) Full Sample HLM Results —Math, Executive Function, and Behavior Skills (Pre-K – 2nd Grade)

	Math				Executive Function				Behavior Skills			
	WJ-III Calculation		WJ-III Applied Problems		Forward Digit Span		Backward Digit Span		SSiS Social Skills		SSiS Problem Behaviors	
Effect	Est[a]	(SE)	Est[a]	(SE)	Est[a]	(SE)	Est[a]	(SE)	Est[a]	(SE)	Est[a]	(SE)
Time x English Proficiency	ns		*** 1>2/3>5, 1>4		ns		ns		***1>4>5, 1>2/3		ns	
Time x Level 1	0.26	(0.65)	5.65***	(0.57)	0.08	(0.04)	0.06	(0.04)	4.81***	(0.95)	-1.87	(0.87)
Time x Level 2 or 3	0.76	(0.39)	1.22**	(0.35)	0.02	(0.02)	0.04	(0.02)	0.88	(0.60)	-0.50	(0.55)
Time x Level 4	0.54	(0.33)	0.59	(0.30)	0.03	(0.02)	0.03	(0.02)	1.33*	(0.52)	-0.43	(0.47)
Time x Level 5	-0.16	(0.61)	0.67	(0.38)	-0.06	(0.04)	-0.05	(0.04)	-4.49***	(0.90)	1.13	(0.83)
Time x IEP[c]	0.96	(0.81)	-0.64	(0.74)	0.08	(0.05)	-0.02	(0.05)	1.56	(1.26)	-1.04	(1.15)
Time x Income[d]	0.33	(0.29)	0.26	(0.26)	-0.02	(0.02)	0.02	(0.02)	0.04	(0.45)	0.32	(0.41)
Time x Provider Type[e]	0.85*	(0.28)	-0.13	(0.25)	0.03	(0.02)	0.00	(0.02)	1.12	(0.44)	-1.17*	(0.42)
Time x Class Size	0.15	(0.10)	-0.01	(0.09)	-0.01	(0.01)	0.00	(0.01)	0.08	(0.16)	0.07	(0.15)
Time x Pre-K CLASS	-0.26	(0.22)	-0.17	(0.27)	-0.02	(0.01)	0.01	(0.01)	-0.09	(0.47)	0.17	(0.45)
Time x CLASS High	-0.04	(0.61)	0.11	(0.50)	0.01	(0.04)	-0.03	(0.04)	1.58	(0.88)	-0.83	(0.82)
Time x CLASS Medium	0.42	(0.58)	0.30	(0.47)	-0.07	(0.03)	0.03	(0.03)	1.46	(0.81)	0.00	(0.76)
Time-sq x Gender[b]	-	-	0.06	(0.03)	-	-	-	-	0.07	(0.06)	0.07	(0.05)
Time-sq x Race			***W>H,B						ns		ns	
Time-sq x Black	-	-	-0.16***	(0.04)	-	-	-	-	-0.07	(0.07)	0.14	(0.06)
Time-sq x Hispanic	-	-	-0.24***	(0.06)	-	-	-	-	0.08	(0.10)	0.12	(0.09)
Time-sq x Multi/Other	-	-	-0.13	(0.07)	-	-	-	-	-0.06	(0.12)	0.07	(0.11)
Time-sq x White	-	-	0.13	(0.07)	-	-	-	-	0.07	(0.12)	-0.07	(0.11)
Time-sq x English Prof			*** 5>1						*** 5>4>1, 2/3>1		ns	
Time-sq x Level 1	-	-	-0.56***	(0.08)	-	-	-	-	-0.65***	(0.13)	0.22	(0.12)
Time-sq x Level 2 or 3	-	-	-0.10	(0.05)	-	-	-	-	-0.10	(0.08)	0.07	(0.08)
Time-sq x Level 4	-	-	-0.05	(0.04)	-	-	-	-	-0.21*	(0.07)	0.07	(0.07)
Time-sq x Level 5	-	-	0.47***	(0.07)	-	-	-	-	0.59***	(0.13)	-0.11	(0.11)
Time-sq x IEP[c]	-	-	0.02	(0.10)	-	-	-	-	-0.06	(0.17)	0.06	(0.16)
Time-sq x Income[d]	-	-	-0.05	(0.04)	-	-	-	-	0.01	(0.06)	-0.06	(0.06)
Time-sq x Provider Type[e]	-	-	0.03	(0.03)	-	-	-	-	-0.10	(0.06)	0.15*	(0.05)
Time-sq x Class Size	-	-	0.00	(0.01)	-	-	-	-	-0.01	(0.02)	-0.01	(0.02)
Time-sq x Pre-K CLASS	-	-	0.00	(0.03)	-	-	-	-	0.01	(0.06)	-0.03	(0.05)
Time-sq x CLASS High	-	-	0.01	(0.07)	-	-	-	-	-0.09	(0.12)	0.04	(0.11)
Time-sq x CLASS Medium	-	-	-0.04	(0.07)	-	-	-	-	-0.09	(0.12)	-0.06	(0.11)

[a] Significance levels are *p< .05, **p< .01, ***p< .001.
[b] Female=0, Male=1.
[c] No IEP=0, IEP=1.
[d] Category Two=0, Category One=1.
[e] Private site=0, Public school site=1.

Figure 1. Growth in Language and Literacy Skills Pre-K – 2nd Grade

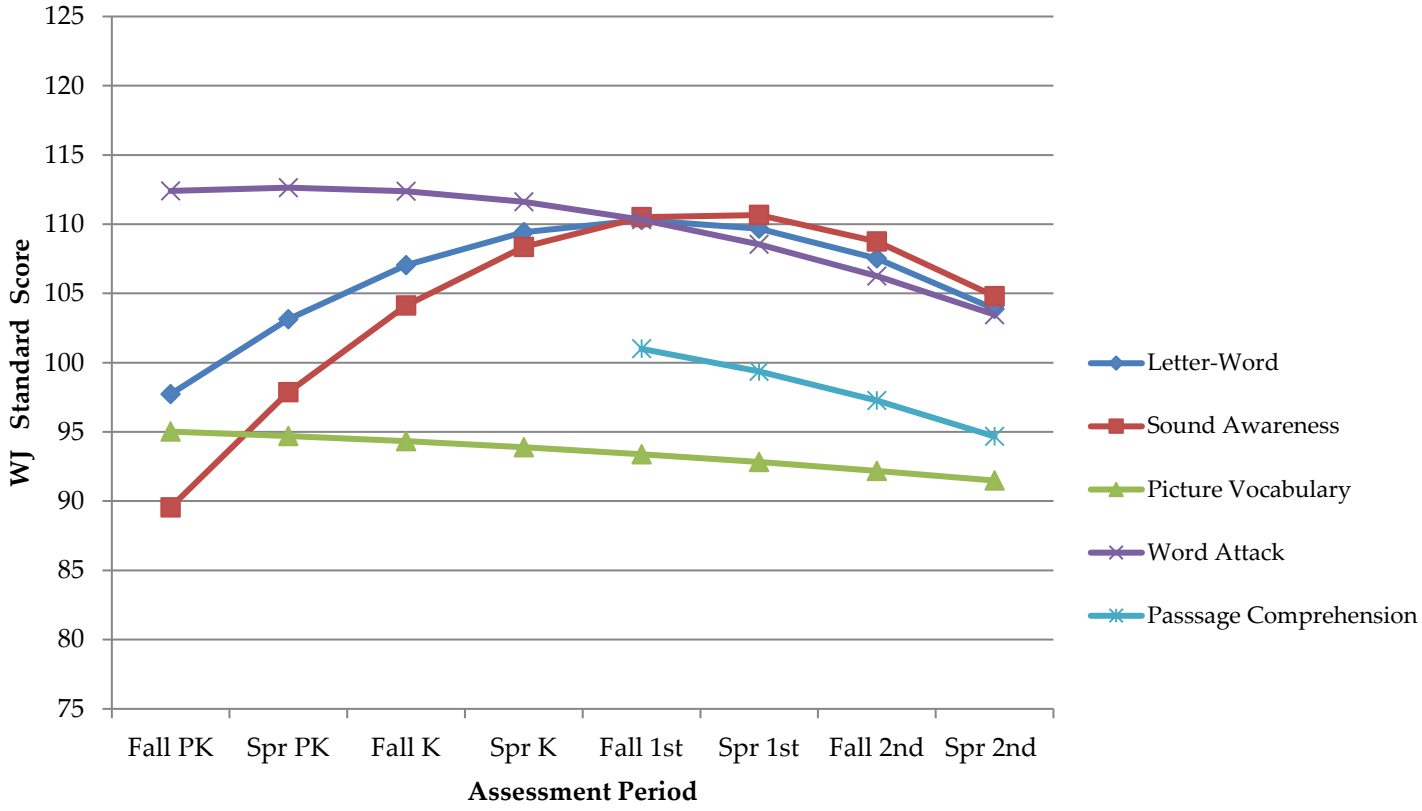


Figure 2. Growth in Math Skills Pre-K – 2nd Grade

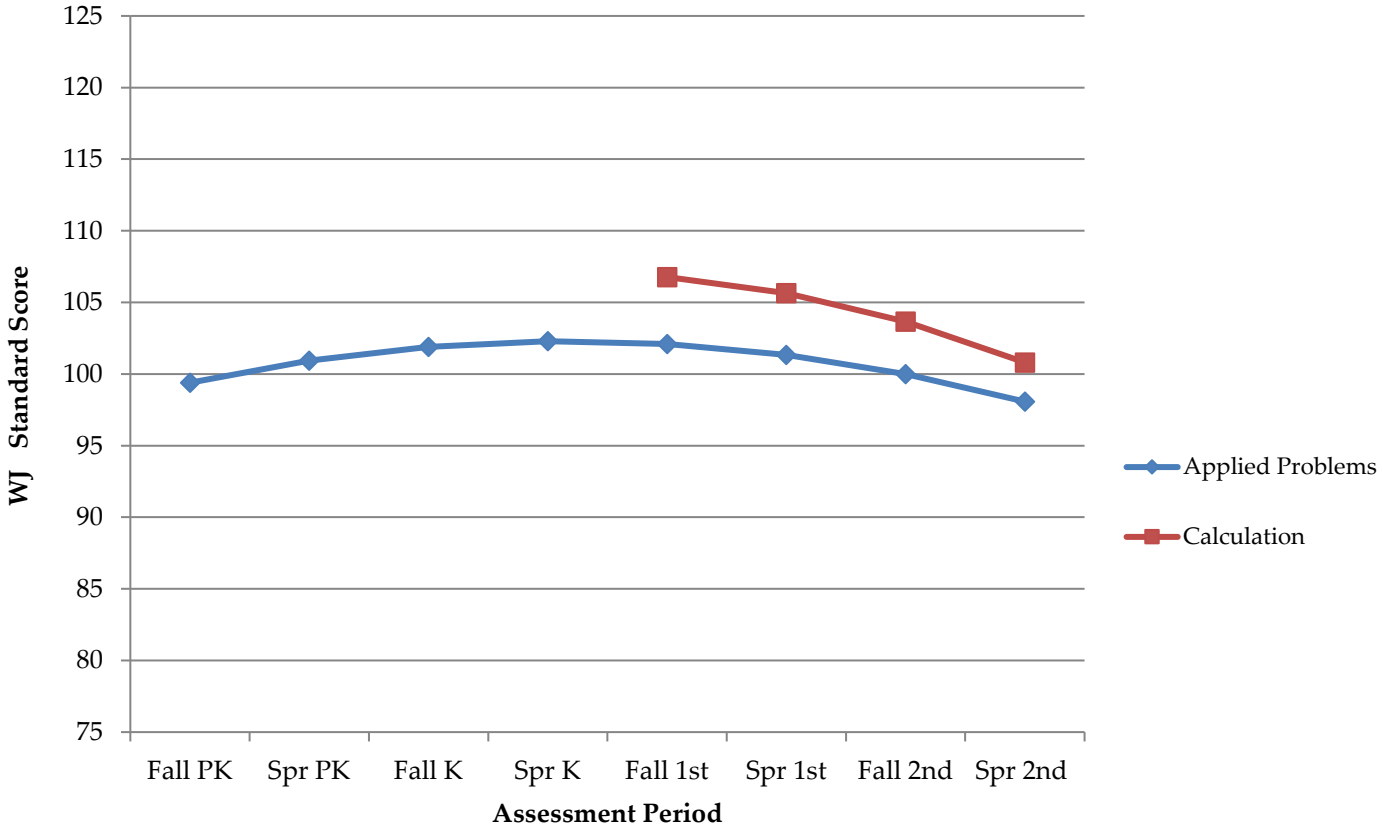


Figure 3. Growth in Behavior Skills Pre-K – 2nd Grade

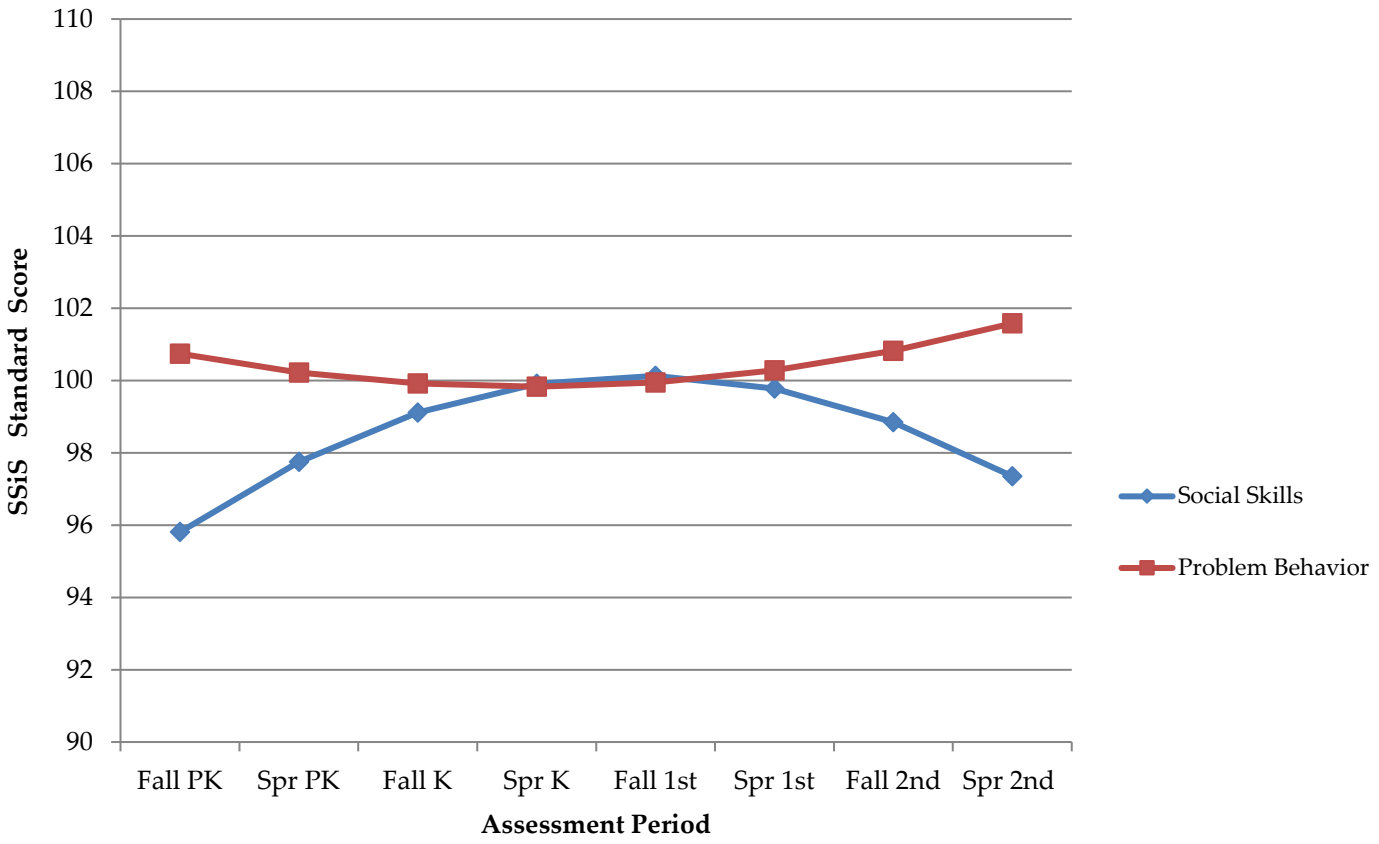


Figure 4. Growth in Executive Function Skills 1st – 2nd Grade

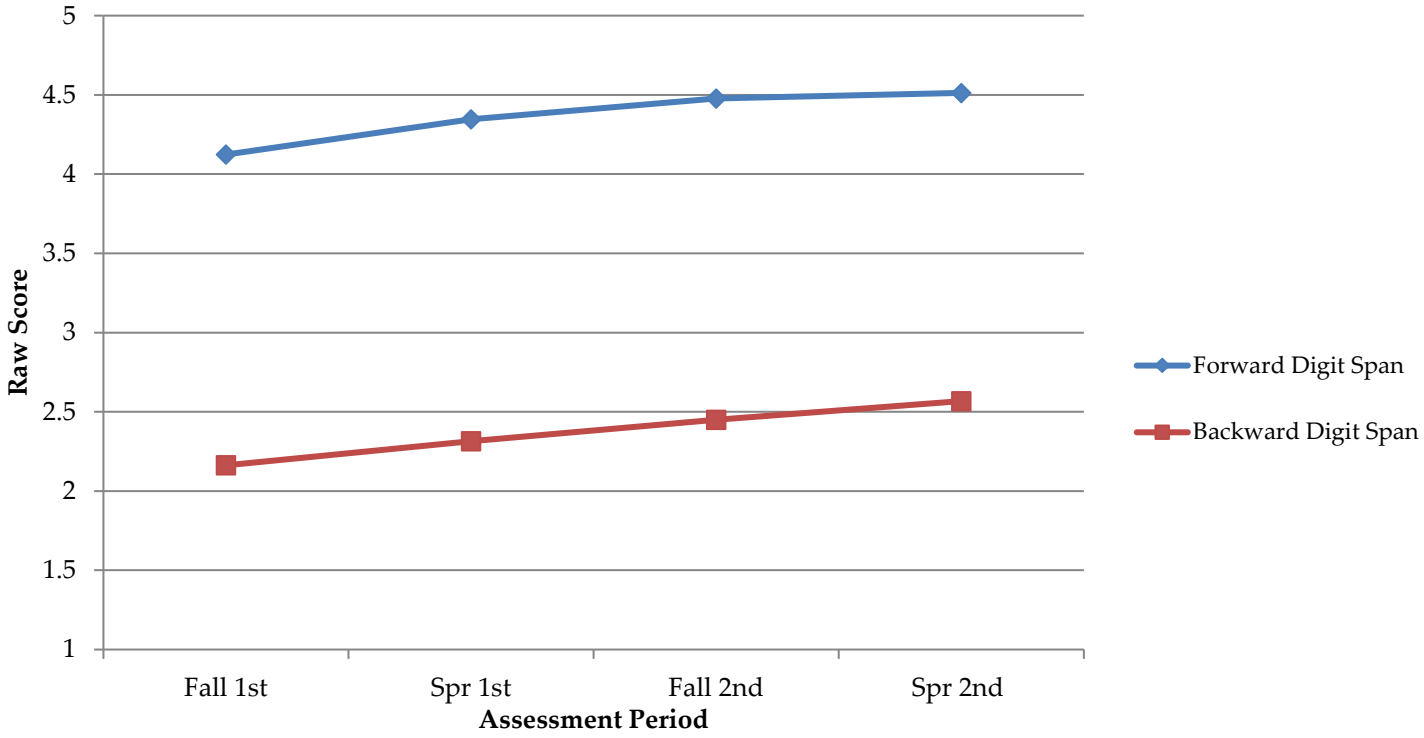


Table 9. Estimated Annual Gains from Pre-K – 2nd Grade (Full Sample)

Measure	Pre-K Gains		Kindergarten Gains		1st Grade Gains		2nd Grade Gains	
	Est[a]	(SE)	Est[a]	(SE)	Est[a]	(SE)	Est[a]	(SE)
Language & Literacy								
WJ-III Picture Vocabulary[b]	-	-	-	-	-	-	-	-
WJ-III Sound Awareness	8.25***	(0.46)	4.41***	(0.25)	0.57*	(0.27)	-3.27***	(0.48)
WJ-III Letter-Word Identification	5.97***	(0.30)	2.56***	(0.18)	-0.86***	(0.19)	-4.27***	(0.32)
WJ-III Word Attack	-0.25	(0.47)	-1.09***	(0.23)	-1.93***	(0.16)	-2.77***	(0.37)
WJ-III Passage Comprehension	-	-	-	-	-3.26***	(0.32)	-4.23***	(0.61)
Math								
WJ-III Calculation	-	-	-	-	-0.93	(0.51)	-2.62***	(0.51)
WJ-III Applied Problems	4.17***	(0.33)	1.90***	(0.17)	-0.38*	(0.18)	-2.66***	(0.34)
Executive Function								
Forward Digit Span	-	-	-	-	0.24***	(0.03)	0.05	(0.03)
Backward Digit Span	-	-	-	-	0.17***	(0.03)	0.13***	(0.03)
Behavior Skills								
SSiS Social Skills	3.87***	(0.56)	1.53***	(0.29)	-0.82**	(0.31)	-3.16***	(0.59)
SSiS Problem Behaviors	-1.13*	(0.51)	-0.41	(0.27)	0.30	(0.29)	1.01	(0.54)

[a] Significance levels are *p< .05, **p< .01, ***p< .001.

[b] Values for this outcome could not be computed based on the model distribution.

Table 10. English Outcome Scores for DLL Subsample (Pre-K – 2nd Grade)

	Pre-K				Kindergarten				1st Grade				2nd Grade			
	Fall		Spring		Fall		Spring		Fall		Spring		Fall		Spring	
Measure	n	Mean (SD) Range	n	Mean (SD) Range	n	Mean (SD) Range	n	Mean (SD) Range	n	Mean (SD) Range	n	Mean (SD) Range	n	Mean (SD) Range	n	Mean (SD) Range
Language & Literacy																
WJ-III Picture Vocabulary[a,b]	129	79.3 (17.3)	125	82.7 (13.2)	125	85.0 (11.4)	122	86.5 (9.4)	122	85.9 (9.7)	121	85.6 (9.9)	113	86.2 (10.3)	118	86.1 (9.4)
WJ-III Sound Awareness[a,b]	129	82.1 (10.9)	126	85.0 (15.8)	124	91.2 (17.5)	124	100.9 (19.6)	122	104.3 (18.4)	121	106.7 (18.3)	113	103.0 (18.6)	118	100.4 (16.5)
WJ-III Letter-Word Identification[a,b]	133	91.2 (14.2)	126	97.8 (12.8)	125	103.1 (13.2)	124	110.3 (12.5)	121	108.0 (12.7)	121	109.2 (11.9)	113	105.8 (12.2)	118	104.0 (11.3)
WJ-III Word Attack[a,b]	23	115.4 (10.3)	64	112.4 (9.2)	120	108.6 (12.4)	124	112.4 (12.1)	122	108.8 (10.8)	121	108.3 (9.9)	113	105.1 (11.3)	118	103.4 (10.9)
WJ-III Passage Comprehension[a,b]	-	-	-	-	-	-	-	-	122	97.9 (13.9)	121	97.0 (10.3)	113	92.9 (9.5)	118	91.6 (10.2)
Math																
WJ-III Calculation[a,b]	-	-	-	-	-	-	-	-	122	106.6 (14.9)	121	107.8 (11.5)	113	101.1 (10.7)	118	97.7 (12.3)
WJ-III Applied Problems[a,b]	131	90.3 (13.7)	125	96.4 (11.3)	125	99.9 (10.4)	124	103.6 (11.8)	122	102.3 (10.8)	121	101.2 (11.0)	113	98.4 (10.7)	118	98.6 (11.0)
Executive Function																
Forward Digit Span[c]	-	-	-	-	-	-	-	-	122	3.9 (0.7)	121	4.1 (0.8)	113	4.0 (0.9)	118	4.3 (0.7)
Backward Digit Span[c]	-	-	-	-	-	-	-	-	122	2.1 (0.6)	121	2.3 (0.6)	113	2.4 (0.6)	118	2.6 (0.7)

[a] Indicates standard score on norm-referenced measure with mean=100, SD=15.

[b] Scores reflect use of updated normative tables (2007).

[c] Possible range=1-8.

Table 11. Spanish Outcome Scores for DLL Subsample (Pre-K – 2nd Grade)

	Pre-K				Kindergarten				1st Grade				2nd Grade			
	Fall		Spring		Fall		Spring		Fall		Spring		Fall		Spring	
Measure	n	Mean (SD) Range	n	Mean (SD) Range	n	Mean (SD) Range	n	Mean (SD) Range	n	Mean (SD) Range	n	Mean (SD) Range	n	Mean (SD) Range	n	Mean (SD) Range
Language & Litteracy																
Bat-III Picture Vocabulary[a,b]	125	74.3 (17.2)	106	72.4 (16.0)	93	70.4 (16.7)	81	68.1 (15.8)	88	66.5 (14.5)	88	64.3 (14.5)	80	65.8 (14.2)	83	64.5 (15.4)
Bat-III Sound Awareness[a,b]	134	74.1 (9.9)	122	77.6 (13.9)	115	76.1 (16.9)	105	83.3 (20.4)	114	90.4 (18.3)	113	90.1 (16.8)	108	89.2 (17.3)	111	85.7 (15.5)
Bat-III Letter-Word Identification[a,b]	135	90.3 (10.7)	123	86.5 (11.7)	118	85.6 (12.9)	107	81.7 (14.1)	116	81.1 (16.8)	115	78.6 (19.4)	108	83.0 (17.5)	115	80.0 (19.5)
Bat-III Word Attack [a,b]	37	107.2 (4.9)	71	101.5 (8.4)	116	95.2 (9.9)	107	90.3 (11.8)	117	87.0 (12.1)	116	85.4 (15.0)	110	84.4 (13.0)	115	81.5 (14.3)
Bat-III Passage Comprehension[a,b]	-	-	-	-	-	-	-	-	103	80.7 (15.1)	103	78.6 (13.2)	106	74.4 (12.5)	111	71.0 (13.8)
Math																
Bat-III Calculation[a,b]	-	-	-	-	-	-	-	-	113	98.7 (19.9)	116	100.2 (17.1)	109	96.7 (14.9)	113	96.5 (15.3)
Bat-III Applied Problems[a,b]	130	87.1 (12.8)	121	91.4 (14.3)	114	91.4 (13.8)	105	94.5 (13.8)	112	94.4 (11.2)	116	91.9 (13.5)	108	88.3 (12.8)	113	86.7 (12.4)
Executive Function																
Forward Digit Span[c]	-	-	-	-	-	-	-	-	117	3.7 (0.8)	117	3.9 (0.8)	110	3.9 (0.8)	115	3.9 (0.8)
Backward Digit Span[c]	-	-	-	-	-	-	-	-	117	2.0 (0.5)	115	2.1 (0.5)	110	2.2 (0.6)	115	2.4 (0.7)

[a] Indicates standard score on norm-referenced measure with mean=100, SD=15.

[b] Scores reflect use of updated normative tables (2007).

[c] Possible range=1-8.

Table 12. DLL Subsample English HLM Results—Language and Literacy Skills (Pre-K – 2nd Grade)

	WJ-III Picture Vocabulary		WJ-III Sound Awareness		WJ-III Letter-Word Identification		WJ-III Word Attack		WJ-III Passage Comprehension	
Effect	Est[a]	(SE)	Est[a]	(SE)	Est[a]	(SE)	Est[a]	(SE)	Est[a]	(SE)
Intercept	79.31	(0.88)	78.36	(1.25)	90.10	(1.44)	108.42	(1.53)	99.38	(1.16)
Child/Family Characteristics										
Gender[b]	0.25	(1.63)	-1.60	(2.20)	-1.03	(2.07)	-1.96	(2.64)	-2.86	(1.92)
English Proficiency										
Level 1	-34.76	(2.95)	-22.88	(4.01)	-13.71	(4.04)	-4.26	(4.60)	-5.99	(3.31)
Level 2 or 3	-15.36	(2.96)	-14.98	(4.00)	-5.98	(3.96)	-5.66	(4.40)	-2.04	(3.43)
Level 4	-9.78	(3.51)	-8.69	(4.72)	-4.12	(4.69)	-4.60	(5.09)	0.31	(4.00)
Family Income[c]	-0.37	(2.13)	1.18	(2.91)	0.74	(2.87)	-1.27	(3.37)	-0.59	(2.43)
Pre-K Characteristics										
Provider Type[d]	-0.70	(1.81)	-4.05	(2.55)	-2.69	(2.91)	-1.36	(3.17)	-3.16	(1.98)
Class Size	-0.28	(0.64)	0.48	(0.90)	0.25	(0.95)	-0.02	(1.22)	0.90	(0.73)
Classroom Quality										
Pre-K CLASS Total	-1.13	(1.74)	2.40	(2.66)	-5.81	(2.50)	-0.41	(2.82)	2.00	(1.52)
K-2 CLASS High	6.78	(2.14)	-4.27	(3.65)	8.70	(2.52)	1.05	(3.90)	0.59	(2.42)
K-2 CLASS Medium	5.58	(1.85)	-9.11	(3.18)	3.15	(2.23)	-2.21	(3.86)	2.81	(2.27)
Time	3.13***	(0.34)	10.53***	(0.60)	8.81***	(0.40)	1.68*	(0.6)	-1.99	(0.88)
Time-squared	-0.33***	(0.04)	-1.02***	(0.07)	-1.01***	(0.05)	-0.35***	(0.07)	-0.11	(0.28)
Time x Gender[b]	0.02	(0.59)	-1.42	(1.05)	0.09	(0.67)	0.00	(1.13)	0.77	(0.52)
Time x English Prof	***1>2/3,4,5		ns		ns		ns		ns	
Time x Level 1	7.66***	(1.03)	-2.69	(1.82)	2.07	(1.18)	-1.92	(1.86)	-0.08	(0.92)
Time x Level 2 or 3	2.50	(1.04)	-1.48	(1.86)	-0.52	(1.19)	-0.90	(1.79)	0.05	(0.95)
Time x Level 4	1.83	(1.23)	1.59	(2.18)	0.07	(1.41)	0.11	(2.04)	0.36	(1.11)
Time x Income[c]	0.32	(0.76)	-0.75	(1.36)	0.25	(0.87)	1.67	(1.37)	-0.04	(0.66)

[a] Significance levels are *p< .05, **p< .01, ***p< .001.
[b] Female=0, Male=1.
[c] Category Two=0, Category One=1.
[d] Private site=0, Public school site=1.

Table 12 (Cont.) DLL Subsample English HLM Results—Language and Literacy Skills (Pre-K – 2nd Grade)

	WJ-III Picture Vocabulary		WJ-III Sound Awareness		WJ-III Letter-Word Identification		WJ-III Word Attack		WJ-III Passage Comprehension	
Effect	Est[a]	(SE)	Est[a]	(SE)	Est[a]	(SE)	Est[a]	(SE)	Est[a]	(SE)
Time x Provider Type[d]	-0.11	(0.62)	1.80	(1.09)	1.08	(0.73)	0.59	(1.21)	-0.55	(0.54)
Time x Class Size	0.23	(0.22)	-0.02	(0.40)	0.26	(0.26)	-0.09	(0.49)	-0.17	(0.19)
Time x Pre-K CLASS	1.26	(0.70)	-0.28	(1.22)	2.74**	(0.83)	-0.34	(1.17)	-0.40	(0.43)
Time x CLASS High	-4.05*	(1.28)	4.19	(2.20)	-3.70	(1.46)	-1.06	(2.15)	-0.47	(1.27)
Time x CLASS Medium	-3.53*	(1.21)	5.12	(2.09)	-1.65	(1.39)	0.43	(2.15)	-2.53*	(1.21)
Time-sq x Gender[b]	0.02	(0.08)	0.24	(0.14)	0.02	(0.09)	0.03	(0.13)	-	-
Time-sq x English Prof	***1<2/3,4,5		*1>4		ns		ns		-	-
Time-sq x Level 1	-0.72***	(0.14)	0.66*	(0.24)	-0.14	(0.15)	0.32	(0.22)	-	-
Time-sq x Level 2 or 3	-0.22	(0.14)	0.49	(0.24)	0.20	(0.15)	0.26	(0.21)	-	-
Time-sq x Level 4	-0.18	(0.16)	0.00	(0.28)	0.12	(0.18)	0.22	(0.25)	-	-
Time-sq x Income[c]	-0.04	(0.10)	0.13	(0.18)	0.05	(0.11)	-0.11	(0.16)	-	-
Time-sq x Provider Type[d]	-0.02	(0.08)	-0.26	(0.14)	-0.21	(0.09)	-0.16	(0.14)	-	-
Time-sq x Class Size	-0.05	(0.03)	0.01	(0.05)	-0.03	(0.03)	0.02	(0.05)	-	-
Time-sq x Pre-K CLASS	-0.15	(0.08)	0.02	(0.14)	-0.25	(0.09)	0.07	(0.13)	-	-
Time-sq x CLASS High	0.52*	(0.18)	-0.63	(0.31)	0.49	(0.20)	0.26	(0.28)	-	-
Time-sq x CLASS Medium	0.42*	(0.17)	-0.61	(0.30)	0.29	(0.20)	0.06	(0.28)	-	-

[a] Significance levels are *p< .05, **p< .01, ***p< .001.
[b] Female=0, Male=1.
[c] Category Two=0, Category One=1.
[d] Private site=0, Public school site=1.

Table 13. DLL Subsample English HLM Results—Math and Executive Function Skills (Pre-K – 2nd Grade)

	Math				Executive Function			
	WJ-III Calculation		WJ-III Applied Problems		Forward Digit Span		Backward Digit Span	
Effect	Est[a]	(SE)	Est[a]	(SE)	Est[a]	(SE)	Est[a]	(SE)
Intercept	108.74	(1.45)	90.54	(0.92)	3.92	(0.09)	2.13	(0.07)
Child/Family Characteristics								
Gender[b]	-1.16	(2.22)	-1.41	(1.77)	0.18	(0.13)	-0.05	(0.10)
English Proficiency								
Level 1	-4.99	(3.86)	-22.28	(3.17)	-0.91	(0.23)	-0.08	(0.18)
Level 2 or 3	-3.28	(4.01)	-8.55	(3.20)	-0.54	(0.24)	0.06	(0.19)
Level 4	2.74	(4.64)	-2.85	(3.78)	-0.58	(0.28)	0.12	(0.22)
Family Income[c]	2.24	(2.84)	1.42	(2.31)	-0.08	(0.17)	0.02	(0.14)
Pre-K Characteristics								
Provider Type[d]	-7.23	(2.41)	-1.40	(1.89)	-0.05	(0.15)	-0.10	(0.12)
Class Size	0.04	(0.85)	0.58	(0.69)	-0.02	(0.05)	-0.02	(0.04)
Classroom Quality								
Pre-K CLASS Total	0.59	(1.86)	0.44	(1.96)	0.07	(0.12)	0.02	(0.09)
K-2 CLASS High	-5.15	(3.29)	0.69	(2.62)	0.05	(0.21)	-0.30	(0.20)
K-2 CLASS Medium	-5.28	(3.09)	-1.95	(2.29)	-0.02	(0.20)	-0.30	(0.19)
Time	0.60	(1.22)	6.08***	(0.43)	0.07	(0.08)	0.15	(0.08)
Time-squared	-1.23*	(0.39)	-0.74***	(0.05)	0.01	(0.03)	0.01	(0.03)
Time x Gender[b]	-0.13	(0.72)	-0.35	(0.72)	0.01	(0.05)	0.09	(0.05)
Time x English Prof	ns		*** 1 > 2/3,4,5		ns		ns	
Time x Level 1	0.74	(1.27)	5.50***	(1.27)	0.15	(0.08)	0.00	(0.08)
Time x Level 2 or 3	0.05	(1.32)	1.55	(1.29)	0.07	(0.09)	-0.04	(0.09)
Time x Level 4	-0.55	(1.53)	-0.15	(1.51)	0.16	(0.10)	-0.14	(0.10)
Time x Income[c]	-0.46	(0.92)	1.12	(0.95)	-0.05	(0.06)	0.01	(0.06)

[a] Significance levels are *p< .05, **p< .01, ***p< .001.
[b] Female=0, Male=1.
[c] Category Two=0, Category One=1.
[d] Private site=0, Public school site=1.

Table 13 (Cont.) DLL Subsample English HLM Results—Math and Executive Function Skills (Pre-K – 2nd Grade)

	Math				Executive Function			
	WJ-III Calculation		WJ-III Applied Problems		Forward Digit Span		Backward Digit Span	
Effect	Est[a]	(SE)	Est[a]	(SE)	Est[a]	(SE)	Est[a]	(SE)
Time x Provider Type[d]	1.65	(0.75)	-0.49	(0.78)	0.04	(0.05)	0.02	(0.05)
Time x Class Size	-0.37	(0.27)	-0.21	(0.28)	-0.02	(0.02)	-0.02	(0.02)
Time x Pre-K CLASS	0.09	(0.59)	-0.20	(0.87)	-0.06	(0.04)	-0.05	(0.04)
Time x CLASS High	3.85	(1.74)	-0.67	(1.57)	-0.01	(0.11)	0.09	(0.11)
Time x CLASS Medium	1.47	(1.66)	-0.78	(1.49)	-0.04	(0.11)	0.07	(0.10)
Time-sq x Gender[b]	-	-	0.13	(0.10)	-	-	-	-
Time-sq x English Prof			** 1>2/3,4,5					
Time-sq x Level 1	-	-	-0.53*	(0.17)	-	-	-	-
Time-sq x Level 2 or 3	-	-	-0.18	(0.17)	-	-	-	-
Time-sq x Level 4	-	-	-0.01	(0.20)	-	-	-	-
Time-sq x Income[c]	-	-	-0.09	(0.12)	-	-	-	-
Time-sq x Provider Type[d]	-	-	0.04	(0.10)	-	-	-	-
Time-sq x Class Size	-	-	0.02	(0.04)	-	-	-	-
Time-sq x Pre-K CLASS	-	-	0.02	(0.10)	-	-	-	-
Time-sq x CLASS High	0.47*	(0.20)	0.07	(0.22)	-	-	-	-
Time-sq x CLASS Medium	0.28	(0.20)	0.20	(0.21)	-	-	-	-

[a] Significance levels are *p< .05, **p< .01, ***p< .001.
[b] Female=0, Male=1.
[c] Category Two=0, Category One=1.
[d] Private site=0, Public school site=1.

Figure 5. DLL Growth in English Language and Literacy Skills Pre-K – 2nd Grade

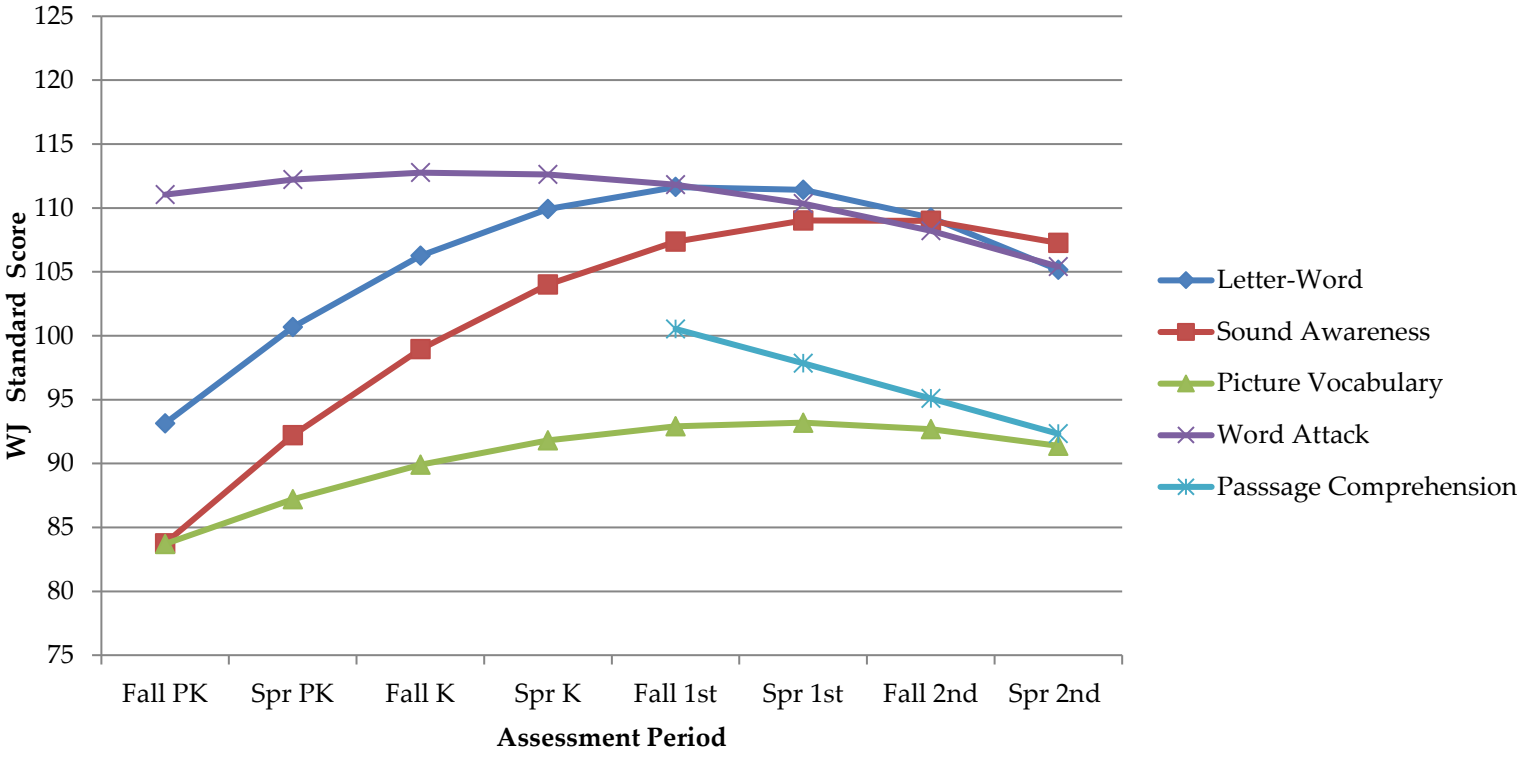


Figure 6. DLL Growth in English Math Skills Pre-K – 2nd Grade

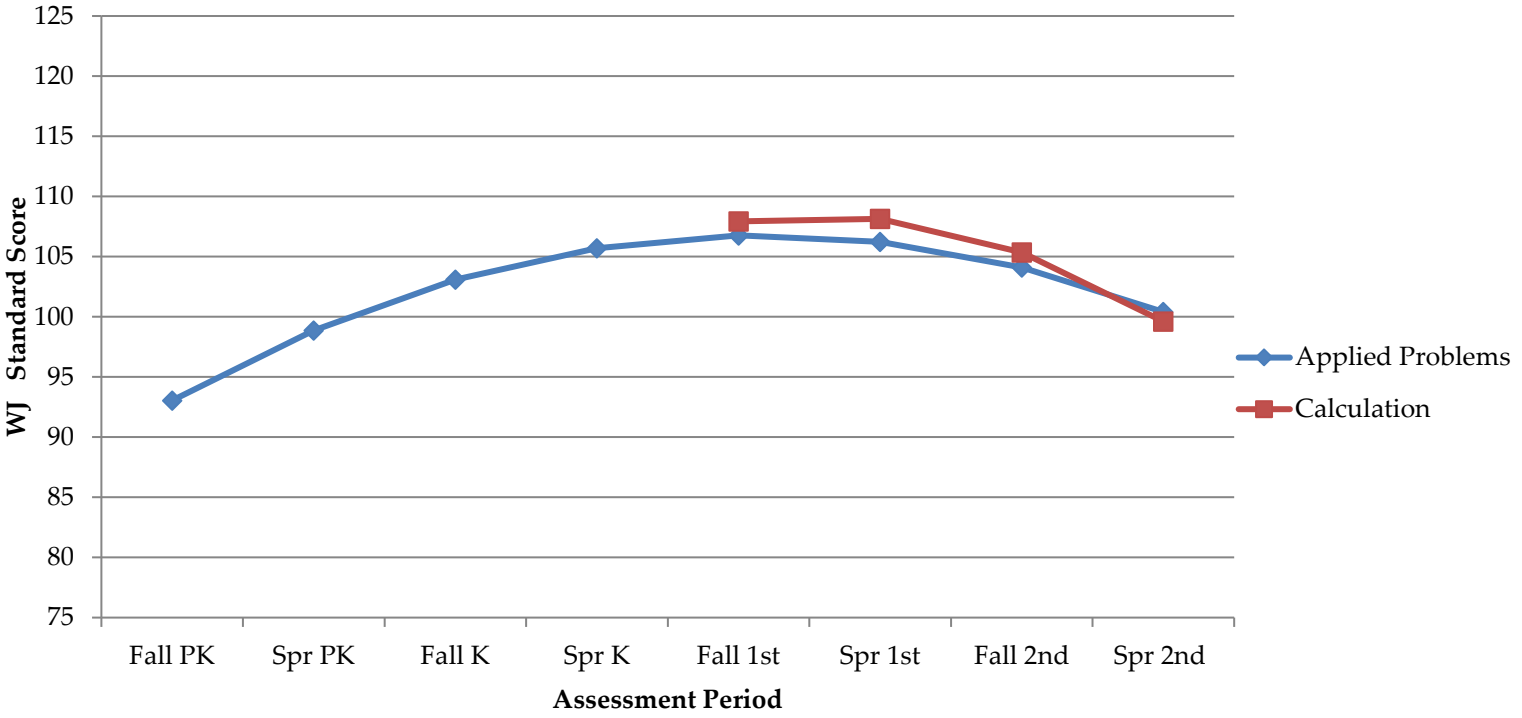


Figure 7. DLL Growth in English Executive Function Skills 1st – 2nd Grade

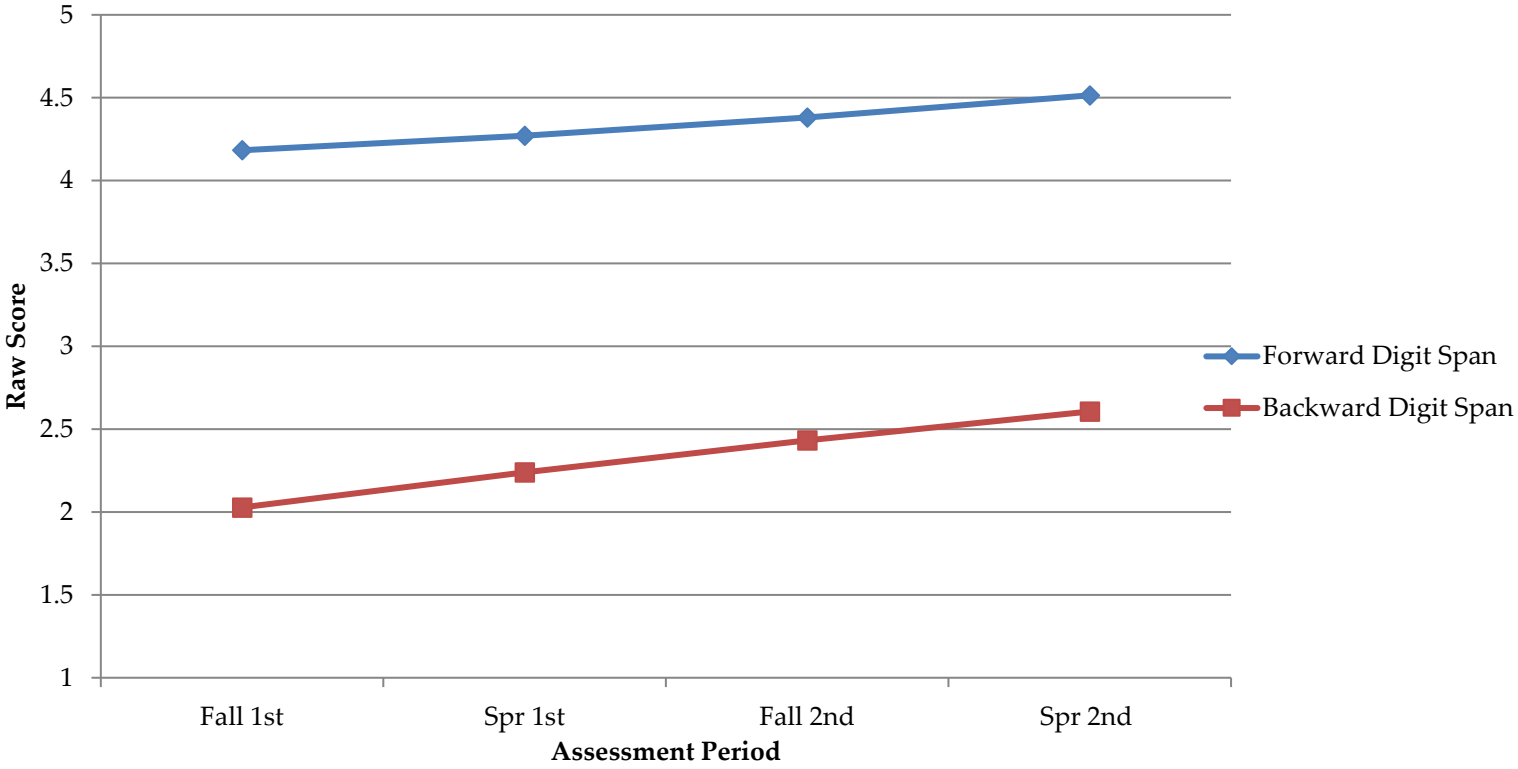


Table 14. Annual Gains in English Outcomes Pre-K – 2nd Grade (DLL Subsample)

	Pre-K Gains		Kindergarten Gains		1st Grade Gains		2nd Grade Gains	
English Outcomes	Est[a]	(SE)	Est[a]	(SE)	Est[a]	(SE)	Est[a]	(SE)
Language & Literacy								
WJ-III Picture Vocabulary	2.80***	(0.30)	1.50***	(0.16)	0.20	(0.14)	-1.11***	(0.27)
WJ-III Sound Awareness	9.51***	(0.53)	5.43***	(0.31)	1.34***	(0.27)	-2.74***	(0.47)
WJ-III Letter-Word Identification	7.80***	(0.36)	3.77***	(0.24)	-0.27	(0.22)	-4.30***	(0.33)
WJ-III Word Attack	1.32	(0.54)	-0.10	(0.32)	-1.51***	(0.23)	-2.93***	(0.39)
WJ-III Passage Comprehension	-	-	-	-	-2.10***	(0.62)	-2.56***	(0.62)
Math								
WJ-III Calculation	-	-	-	-	-0.63	(0.86)	-5.54***	(0.86)
WJ-III Applied Problems	5.34***	(0.38)	2.39***	(0.22)	-0.57*	(0.20)	-3.52***	(0.34)
Executive Function								
Forward Digit Span	-	-	-	-	0.08	(0.06)	0.12	(0.06)
Backward Digit Span	-	-	-	-	0.16*	(0.06)	0.18*	(0.06)

[a] Significance levels are *p< .05, **p< .01, ***p< .001.

Table 15. DLL Subsample Spanish HLM Results—Language and Literacy Skills (Pre-K – 2nd Grade)

	Bat-III Picture Vocabulary		Bat-III Sound Awareness		Bat-III Letter-Word Identification		Bat-III Word Attack		Bat-III Passage Comprehension	
Effect	Est[a]	(SE)	Est[a]	(SE)	Est[a]	(SE)	Est[a]	(SE)	Est[a]	(SE)
Intercept	73.28	(1.17)	71.85	(1.13)	89.40	(0.96)	106.07	(1.19)	79.26	(1.53)
Child/Family Characteristics										
Gender[b]	-1.47	(2.41)	-1.17	(2.33)	-2.84	(1.98)	-4.02	(2.33)	-1.62	(2.26)
Spanish Proficiency										
Level 1	-29.38	(3.26)	-3.00	(3.14)	-7.90	(2.67)	2.18	(3.02)	-19.07	(3.03)
Level 2 or 3	-19.46	(3.37)	-5.16	(3.37)	-4.62	(2.88)	0.31	(3.18)	-14.35	(3.25)
Level 4	-9.95	(3.40)	-5.53	(3.41)	-0.25	(2.92)	1.88	(3.48)	-4.51	(3.21)
Family Income[c]	6.17	(2.88)	-2.42	(2.73)	-4.96	(2.29)	-2.85	(2.59)	0.62	(2.65)
Pre-K Characteristics										
Provider Type[d]	2.94	(2.41)	-0.61	(2.34)	-2.60	(1.99)	-1.20	(2.43)	-1.40	(2.49)
Class Size	1.41	(0.87)	0.26	(0.85)	0.83	(0.72)	1.43	(0.86)	-0.21	(0.82)
Classroom Quality										
Pre-K CLASS Total	3.25	(2.47)	2.63	(2.76)	0.19	(2.28)	-2.17	(2.38)	0.66	(1.94)
K-2 CLASS High	-2.52	(3.16)	-2.91	(4.11)	3.49	(3.33)	2.54	(3.86)	-4.63	(3.40)
K-2 CLASS Medium	-1.67	(2.72)	-5.25	(3.46)	0.74	(2.83)	2.52	(3.66)	-2.35	(3.26)
Time	-4.84***	(0.52)	6.85***	(0.69)	-3.52***	(0.56)	-6.31***	(0.63)	-2.27	(1.32)
Time-squared	0.43***	(0.07)	-0.63***	(0.09)	0.30***	(0.07)	0.41***	(0.07)	-0.09	(0.41)
Time x Gender[b]	0.25	(0.91)	-2.80	(1.26)	-0.31	(1.02)	2.06	(1.16)	0.07	(0.80)
Time x Spanish Prof	ns		***1<4,5		*1 < 5		***1<4,5		ns	
Time x Level 1	0.27	(1.29)	-6.30**	(1.70)	-4.66**	(1.37)	-5.71**	(1.51)	0.39	(1.07)
Time x Level 2 or 3	-1.82	(1.19)	-4.45	(1.80)	-3.81	(1.46)	-3.28	(1.61)	0.07	(1.14)
Time x Level 4	-2.24	(1.18)	0.54	(1.81)	-1.26	(1.48)	-0.62	(1.73)	0.01	(1.13)
Time x Income[c]	-1.30	(1.09)	0.20	(1.39)	-0.31	(1.13)	0.68	(1.15)	-0.19	(0.63)

[a] Significance levels are *p< .05, **p< .01, ***p< .001.
[b] Female=0, Male=1.
[c] Category Two=0, Category One=1.
[d] Private site=0, Public school site=1.

Table 15 (Cont.) DLL Subsample Spanish HLM Results—Language and Literacy Skills (Pre-K – 2nd Grade)

	Bat-III Picture Vocabulary		Bat-III Sound Awareness		Bat-III Letter-Word Identification		Bat-III Word Attack		Bat-III Passage Comprehension	
Effect	Est[a]	(SE)	Est[a]	(SE)	Est[a]	(SE)	Est[a]	(SE)	Est[a]	(SE)
Time x Provider Type[d]	0.10	(0.32)	0.26	(0.45)	-0.02	(0.36)	-0.62	(0.43)	0.01	(0.28)
Time x Class Size	-0.84	(1.88)	1.31	(2.57)	-1.26	(2.05)	-2.97	(2.21)	2.32	(1.79)
Time x Pre-K CLASS	-0.68	(0.91)	-0.25	(1.27)	1.41	(1.03)	-0.51	(1.21)	-1.00	(0.81)
Time x CLASS High	-0.31	(1.78)	-0.91	(2.44)	-0.29	(1.94)	-3.16	(2.17)	0.94	(1.81)
Time x CLASS Medium	-0.31	(1.78)	0.44	(0.17)	0.17	(0.13)	-0.14	(0.14)	1.11	(1.81)
Time-sq x Gender[b]	-0.15	(0.12)	0.41*	(0.17)	-3.65	(1.45)	-0.16	(0.14)	-	-
Time-sq x Spanish Prof	ns		*1>4		ns		ns		-	-
Time-sq x Level 1	0.01	(0.17)	0.56	(0.23)	0.33	(0.18)	0.43	(0.18)	-	-
Time-sq x Level 2 or 3	0.17	(0.16)	0.57	(0.24)	0.27	(0.19)	0.24	(0.19)	-	-
Time-sq x Level 4	0.20	(0.15)	0.00	(0.24)	0.01	(0.19)	-0.07	(0.21)	-	-
Time-sq x Income[c]	0.14	(0.15)	0.05	(0.20)	-0.08	(0.15)	-0.01	(0.15)	-	-
Time-sq x Provider Type[d]	0.08	(0.12)	-0.01	(0.17)	-0.19	(0.13)	0.01	(0.14)	-	-
Time-sq x Class Size	-0.05	(0.04)	-0.04	(0.06)	0.00	(0.05)	0.09	(0.05)	-	-
Time-sq x Pre-K CLASS	0.11	(0.13)	-0.06	(0.17)	0.06	(0.13)	-0.01	(0.13)	-	-
Time-sq x CLASS High	0.16	(0.26)	-0.16	(0.37)	0.20	(0.29)	0.55	(0.29)	-	-
Time-sq x CLASS Medium	0.20	(0.26)	0.31	(0.37)	0.06	(0.29)	0.53	(0.29)	-	-

[a] Significance levels are *p< .05, **p< .01, ***p< .001.
[b] Female=0, Male=1.
[c] Category Two=0, Category One=1.
[d] Private site=0, Public school site=1.

Table 16. DLL Subsample Spanish HLM Results—Math and Executive Function Skills (Pre-K – 2nd Grade)

	Math				Executive Function			
	Bat-III Calculation		Bat-III Applied Problems		Forward Digit Span		Backward Digit Span	
Effect	Est[a]	(SE)	Est[a]	(SE)	Est[a]	(SE)	Est[a]	(SE)
Intercept	100.08	(2.06)	86.74	(0.97)	3.79	(0.10)	1.95	(0.06)
Child/Family Characteristics								
Gender[b]	-7.80	(4.34)	-2.56	(2.03)	0.09	(0.15)	-0.10	(0.10)
Spanish Proficiency								
Level 1	-7.81	(4.66)	-19.74	(2.72)	-0.49	(0.21)	-0.09	(0.13)
Level 2 or 3	-0.11	(4.67)	-9.57	(2.93)	-0.31	(0.22)	-0.14	(0.14)
Level 4	-6.33	(3.78)	-5.88	(2.97)	-0.08	(0.22)	0.09	(0.14)
Family Income[c]	3.58	(2.61)	-0.36	(2.36)	-0.37	(0.18)	0.07	(0.11)
Pre-K Characteristics								
Provider Type[d]	0.11	(1.14)	-0.33	(2.03)	-0.02	(0.17)	-0.06	(0.10)
Class Size	1.74	(4.69)	0.75	(0.73)	0.00	(0.06)	-0.04	(0.03)
Classroom Quality								
Pre-K CLASS Total	-7.18	(3.37)	4.32	(2.10)	0.13	(0.13)	-0.04	(0.08)
K-2 CLASS High	1.31	(4.48)	-5.51	(2.75)	0.26	(0.20)	-0.14	(0.17)
K-2 CLASS Medium	1.02	(4.46)	-9.37	(2.31)	0.23	(0.19)	-0.08	(0.16)
Time	0.46	(1.84)	3.87***	(0.44)	0.14	(0.08)	0.13	(0.07)
Time-squared	-0.46	(0.58)	-0.56***	(0.06)	-0.02	(0.02)	0.00	(0.02)
Time x Gender[b]	1.12	(1.12)	-0.57	(0.79)	-0.02	(0.05)	0.06	(0.04)
Time x Spanish Prof	ns		ns		ns		ns	
Time x Level 1	0.87	(1.50)	-0.33	(1.06)	0.02	(0.06)	-0.13	(0.06)
Time x Level 2 or 3	0.71	(1.59)	0.16	(1.13)	0.06	(0.07)	-0.05	(0.06)
Time x Level 4	0.89	(1.64)	1.04	(1.15)	0.13	(0.07)	-0.07	(0.06)
Time x Income[c]	2.54	(1.29)	1.87	(0.95)	0.09	(0.05)	0.03	(0.05)

[a] Significance levels are *p< .05, **p< .01, ***p< .001.
[b] Female=0, Male=1.
[c] Category Two=0, Category One=1.
[d] Private site=0, Public school site=1.

Table 16 (Cont.) DLL Subsample Spanish HLM Results—Math and Executive Function Skills (Pre-K – 2nd Grade)

	Math				Executive Function			
	Bat-III Calculation		Bat-III Applied Problems		Forward Digit Span		Backward Digit Span	
Effect	Est[a]	(SE)	Est[a]	(SE)	Est[a]	(SE)	Est[a]	(SE)
Time x Provider Type	1.40	(1.13)	-0.78	(0.80)	0.01	(0.05)	0.00	(0.04)
Time x Class Size	-0.18	(0.38)	-0.08	(0.28)	-0.03	(0.02)	-0.01	(0.01)
Time x Pre-K CLASS	-0.29	(0.89)	-2.68*	(0.91)	-0.10*	(0.04)	0.02	(0.03)
Time x CLASS High	1.10	(2.48)	2.35	(1.67)	-0.18	(0.11)	0.18	(0.09)
Time x CLASS Medium	2.23	(2.51)	4.13	(1.57)	-0.08	(0.11)	0.18	(0.09)
Time-sq x Gender[b]	-	-	0.09	(0.11)	-	-	-	-
Time-sq x Spanish Prof			ns					
Time-sq x Level 1	-	-	0.09	(0.15)	-	-	-	-
Time-sq x Level 2 or 3	-	-	-0.06	(0.15)	-	-	-	-
Time-sq x Level 4	-	-	-0.07	(0.16)	-	-	-	-
Time-sq x Income[c]	-	-	-0.21	(0.13)	-	-	-	-
Time-sq x Provider Type[d]	-	-	0.08	(0.11)	-	-	-	-
Time-sq x Class Size	-	-	0.00	(0.04)	-	-	-	-
Time-sq x Pre-K CLASS	-	-	0.32*	(0.11)	-	-	-	-
Time-sq x CLASS High	-	-	-0.21	(0.23)	-	-	-	-
Time-sq x CLASS Medium	-	-	-0.38	(0.23)	-	-	-	-

[a] Significance levels are *p< .05, **p< .01, ***p< .001.
[b] Female=0, Male=1.
[c] Category Two=0, Category One=1.
[d] Private site=0, Public school site=1.

Figure 8. DLL Growth in Spanish Language and Literacy Skills Pre-K – 2nd Grade

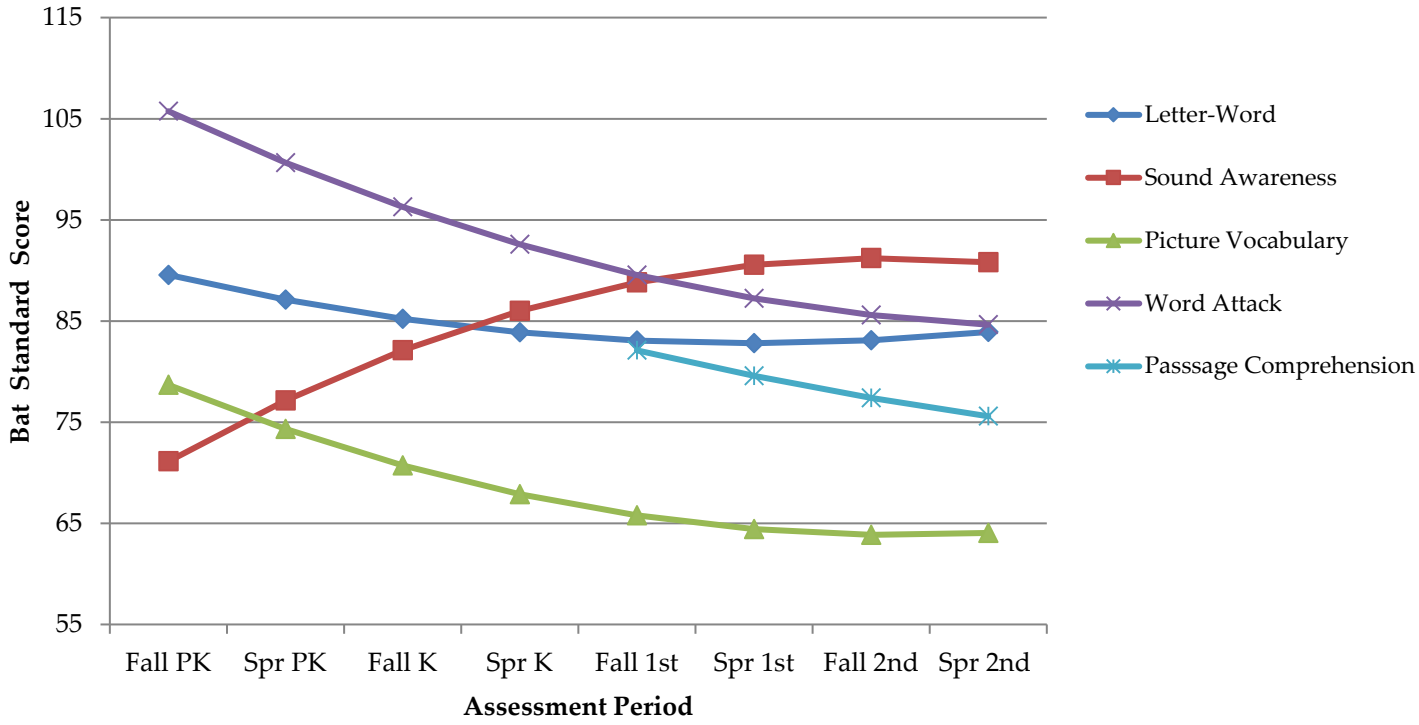


Figure 9. DLL Growth in Spanish Math Skills Pre-K – 2nd Grade

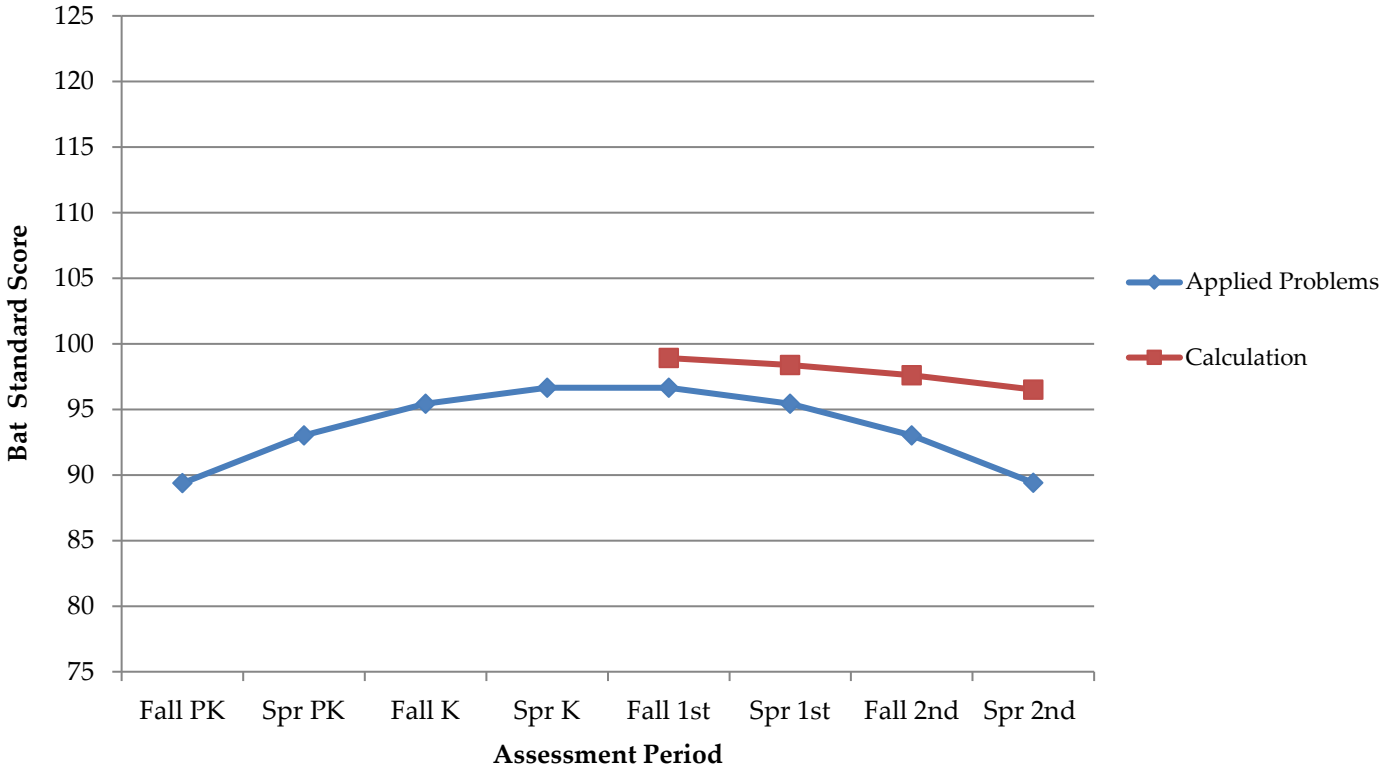


Figure 10. DLL Growth in Spanish Executive Function Skills 1st – 2nd Grade

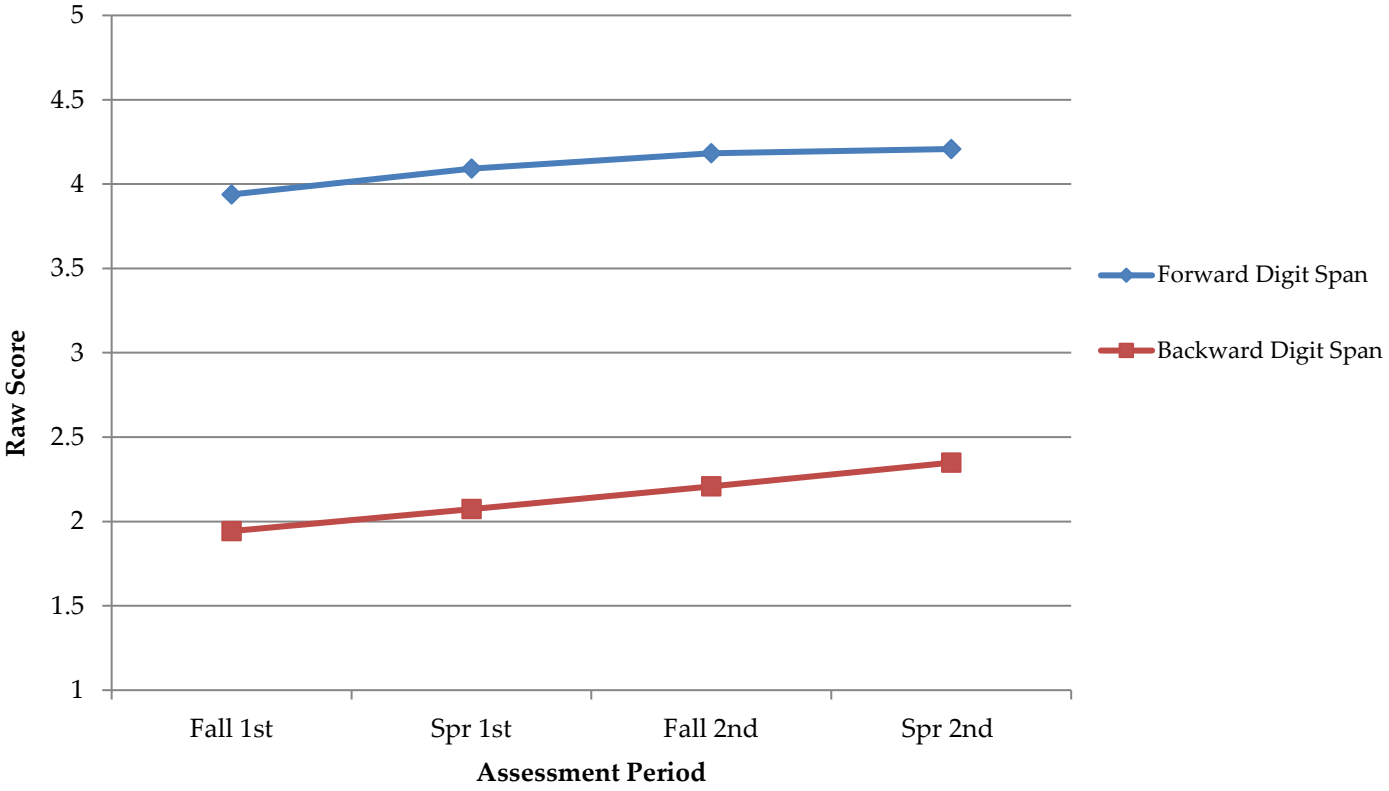


Table 17. Annual Gains in Spanish Outcomes Pre-K – 2nd Grade (DLL Subsample)

	Pre-K Gains		Kindergarten Gains		1st Grade Gains		2nd Grade Gains	
Spanish Outcomes	Est[a]	(SE)	Est[a]	(SE)	Est[a]	(SE)	Est[a]	(SE)
Language & Literacy								
Bat-III Picture Vocabulary	-4.41***	(0.46)	-2.71***	(0.27)	-1.01***	(0.26)	0.70	(0.45)
Bat-III Sound Awareness	6.22***	(0.62)	3.69***	(0.35)	1.16***	(0.31)	-1.36	(0.56)
Bat-III Letter-Word Identification	-3.22***	(0.50)	-2.03***	(0.30)	-0.83*	(0.27)	0.36	(0.45)
Bat-III Word Attack	-5.90***	(0.56)	-4.27***	(0.32)	-2.64***	(0.22)	-1.01	(0.41)
Bat-III Passage Comprehension					-2.36	(0.93)	-2.72*	(0.90)
Math								
Bat-III Calculation	-	-	-	-	0.00	(1.30)	-1.82	(1.28)
Bat-III Applied Problems	3.31***	(0.39)	1.08***	(0.20)	-1.15***	(0.18)	-3.38***	(0.35)
Executive Function								
Forward Digit Span	-	-	-	-	0.12	(0.05)	0.03	(0.05)
Backward Digit Span	-	-	-	-	0.13	(0.05)	0.12	(0.05)

[a] Significance levels are *p< .05, **p< .01, ***p< .001.

Figure 11. Growth in WJ-III Picture Vocabulary by English Proficiency

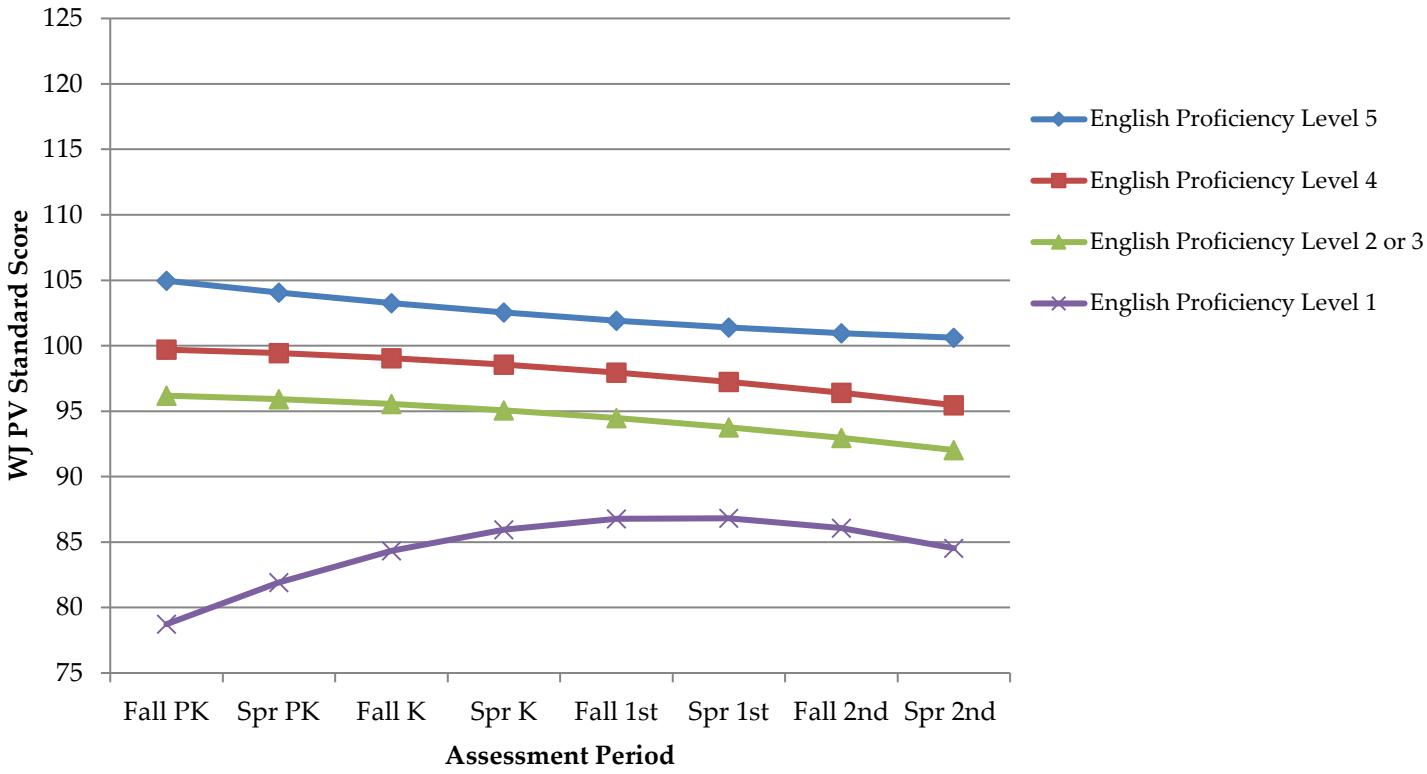


Figure 12. Growth in WJ-III Sound Awareness by English Proficiency

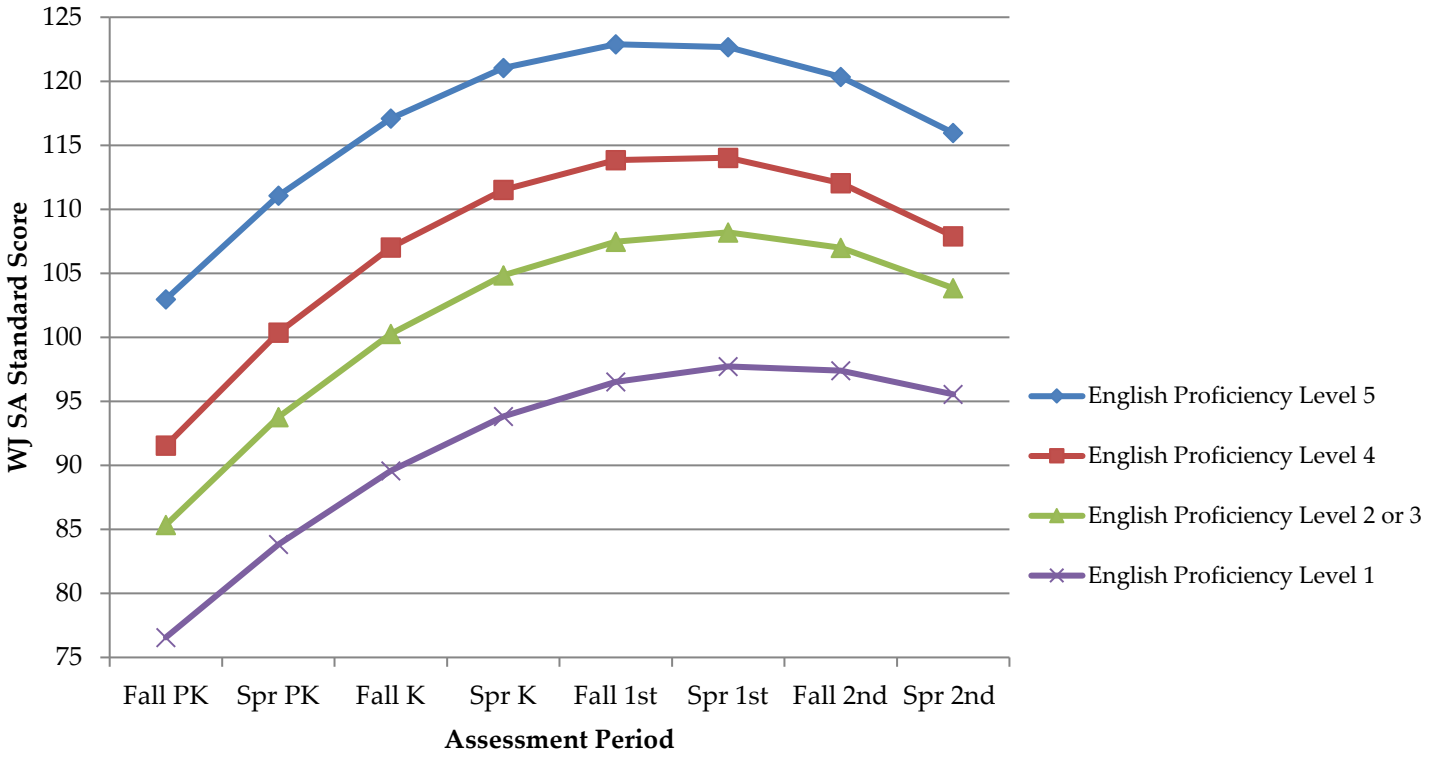


Figure 13. Growth in WJ-III Letter-Word Identification by English Proficiency

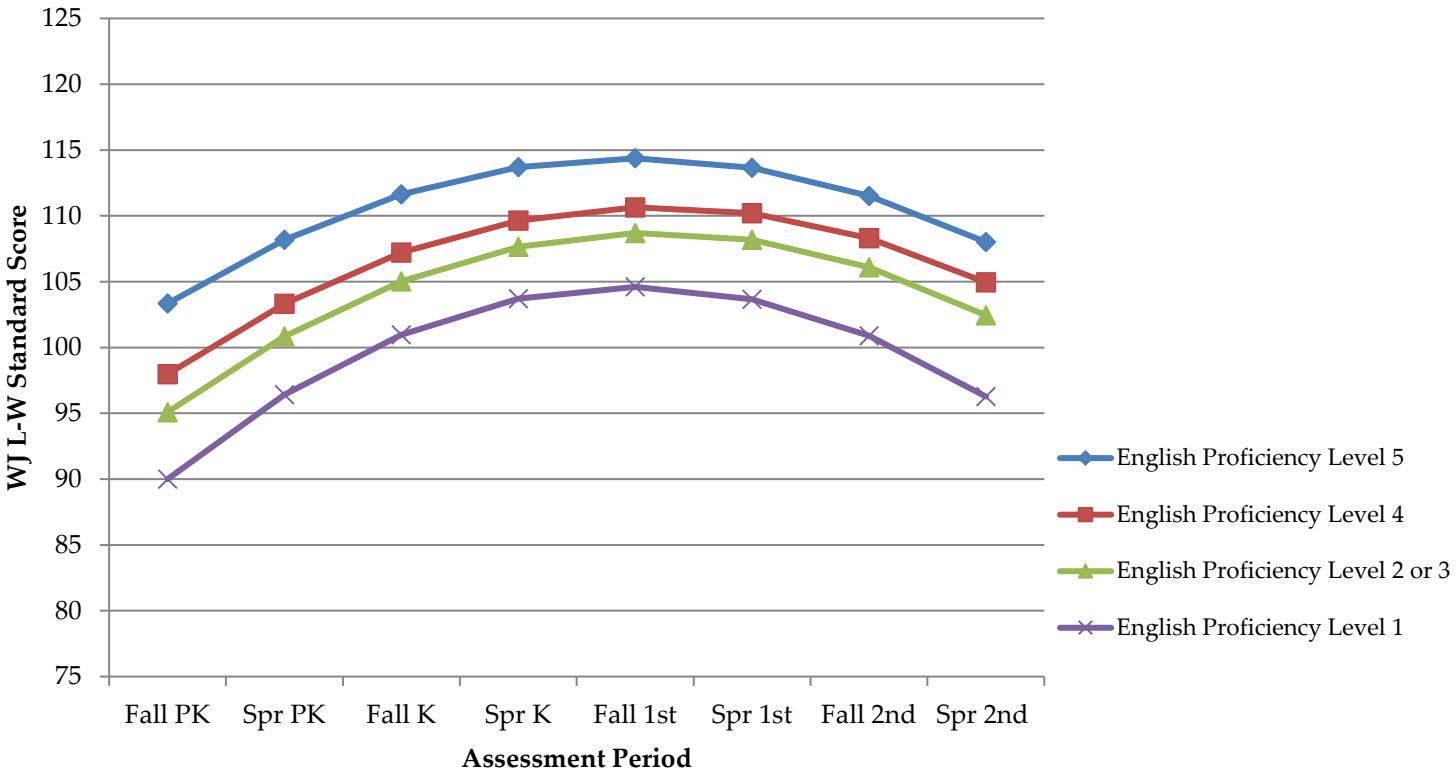


Figure 14. Growth in WJ-III Passage Comprehension by English Proficiency

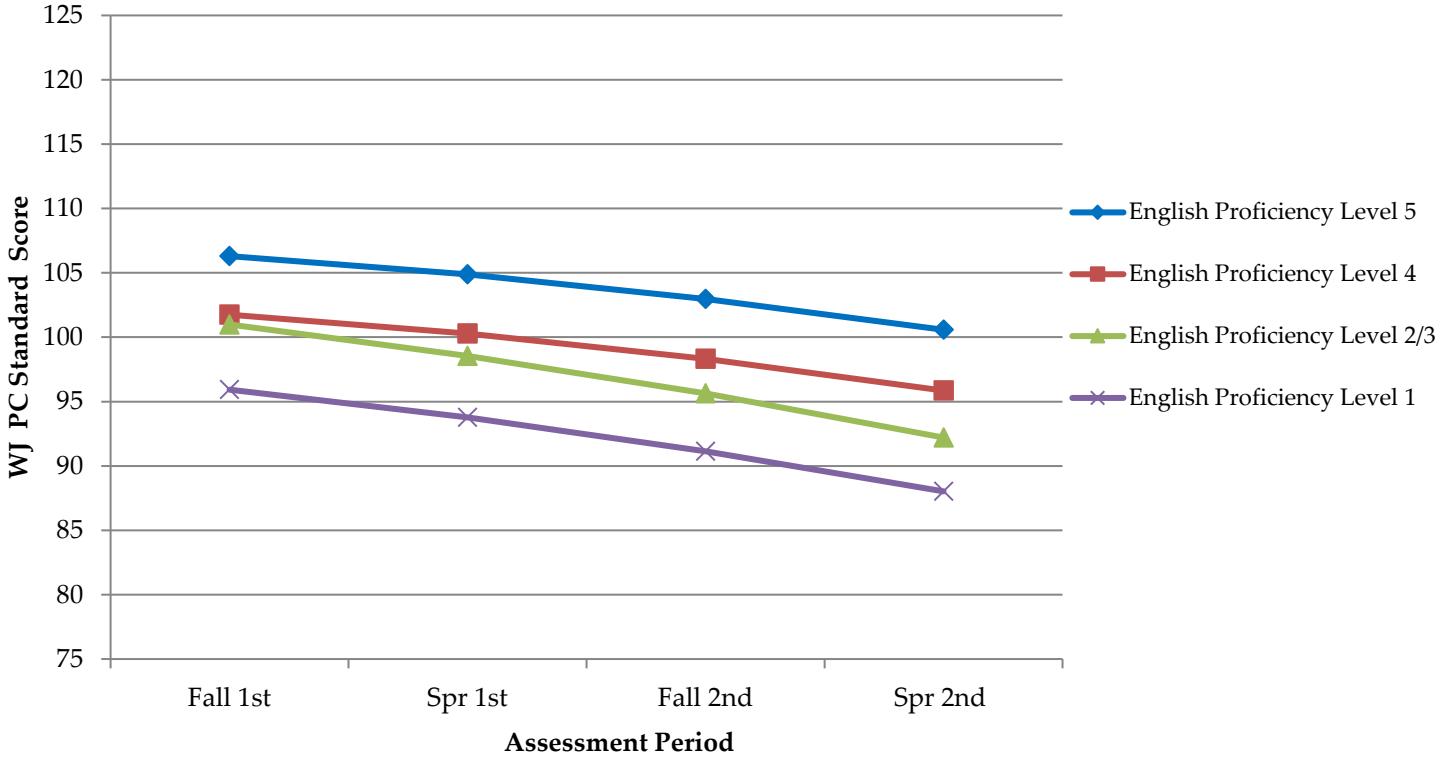


Figure 15. Growth in WJ-III Applied Problems by English Proficiency

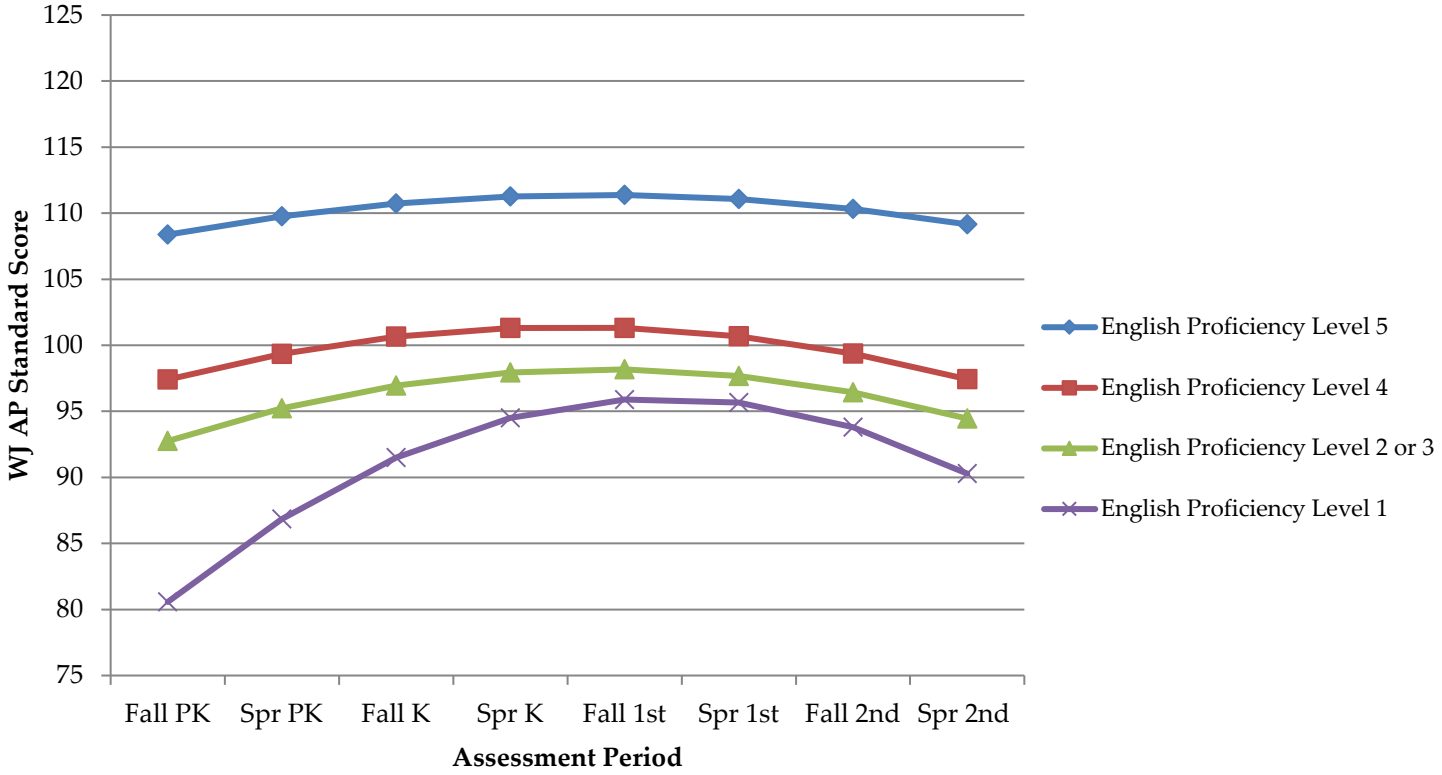


Figure 16. Growth in SSiS Social Skills by English Proficiency

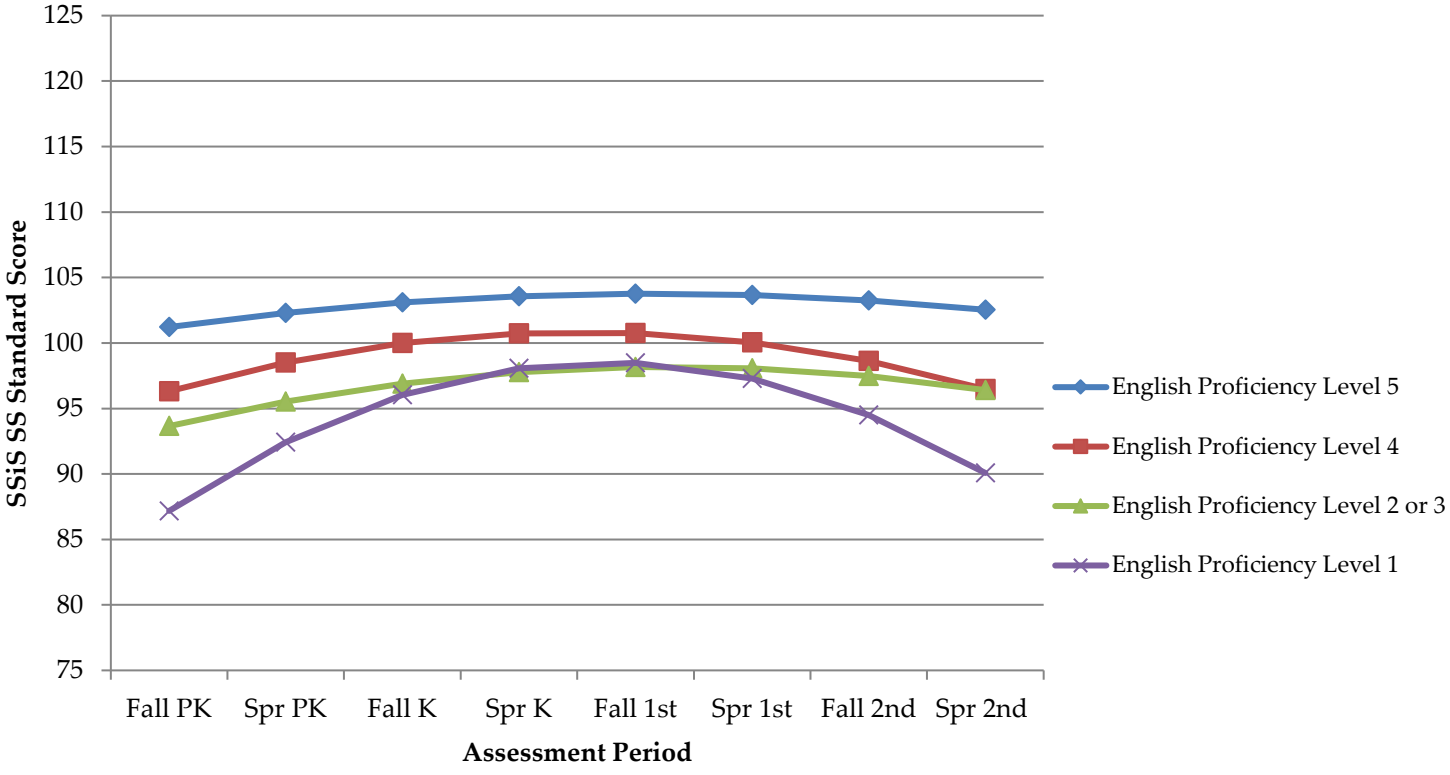


Figure 17. Growth in WJ-III Picture Vocabulary by Race/Ethnicity

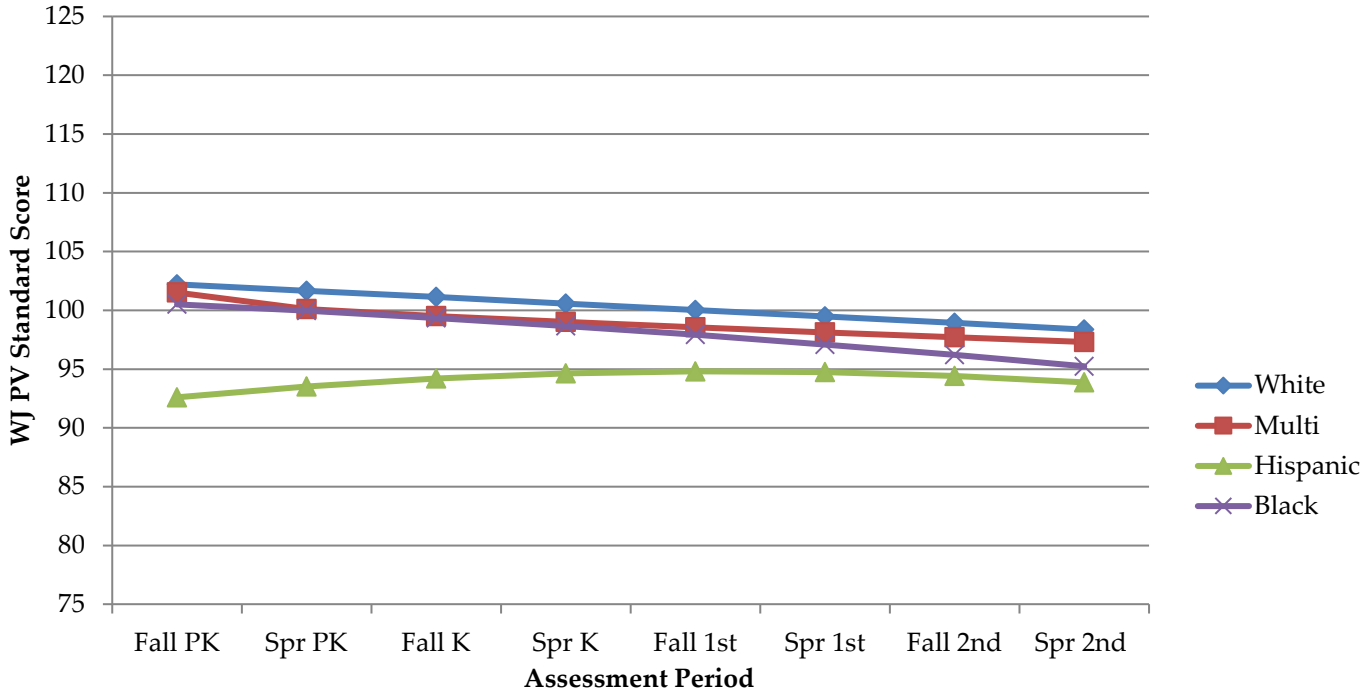


Figure 18. Growth in WJ-III Letter-Word Identification by Race/Ethnicity

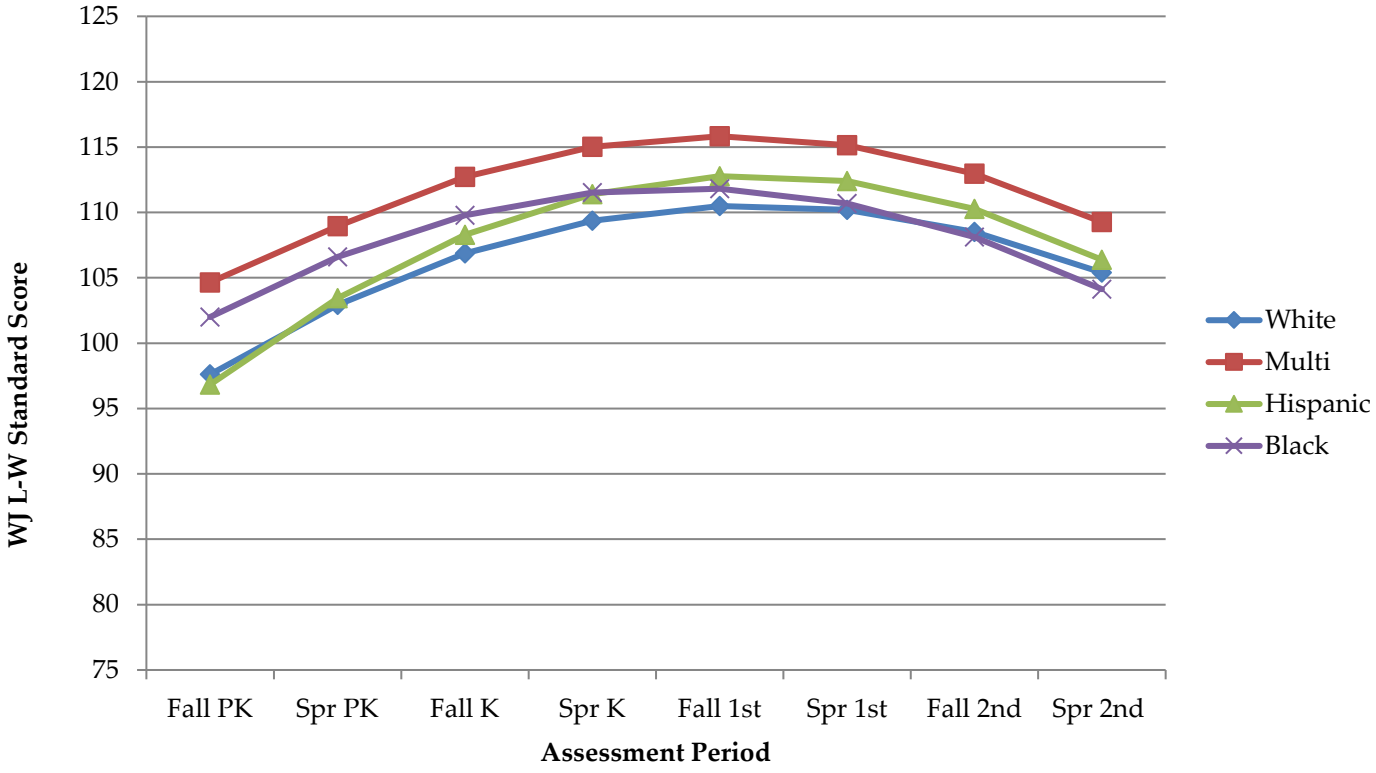


Figure 19. Growth in WJ-III Word Attack by Race/Ethnicity

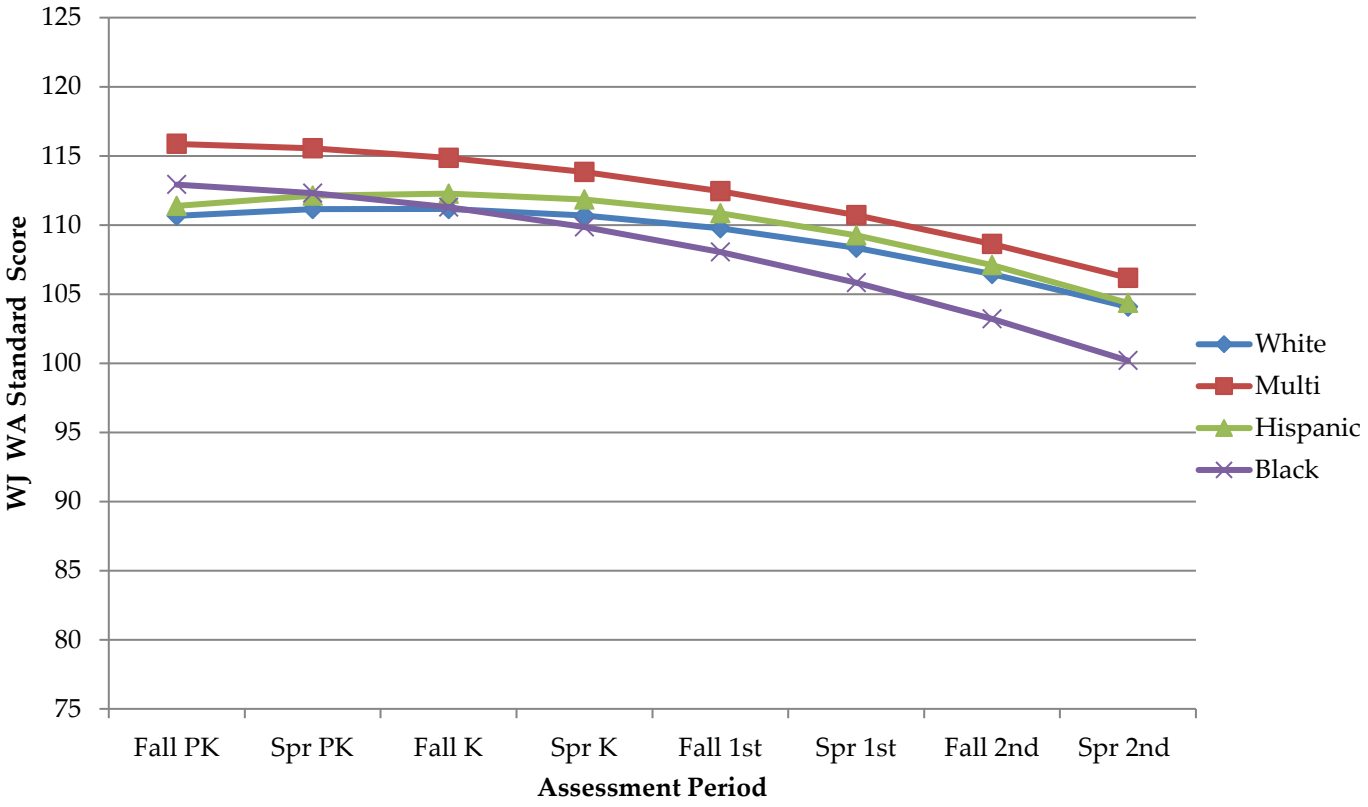


Figure 20. Growth in WJ-III Passage Comprehension by Race/Ethnicity

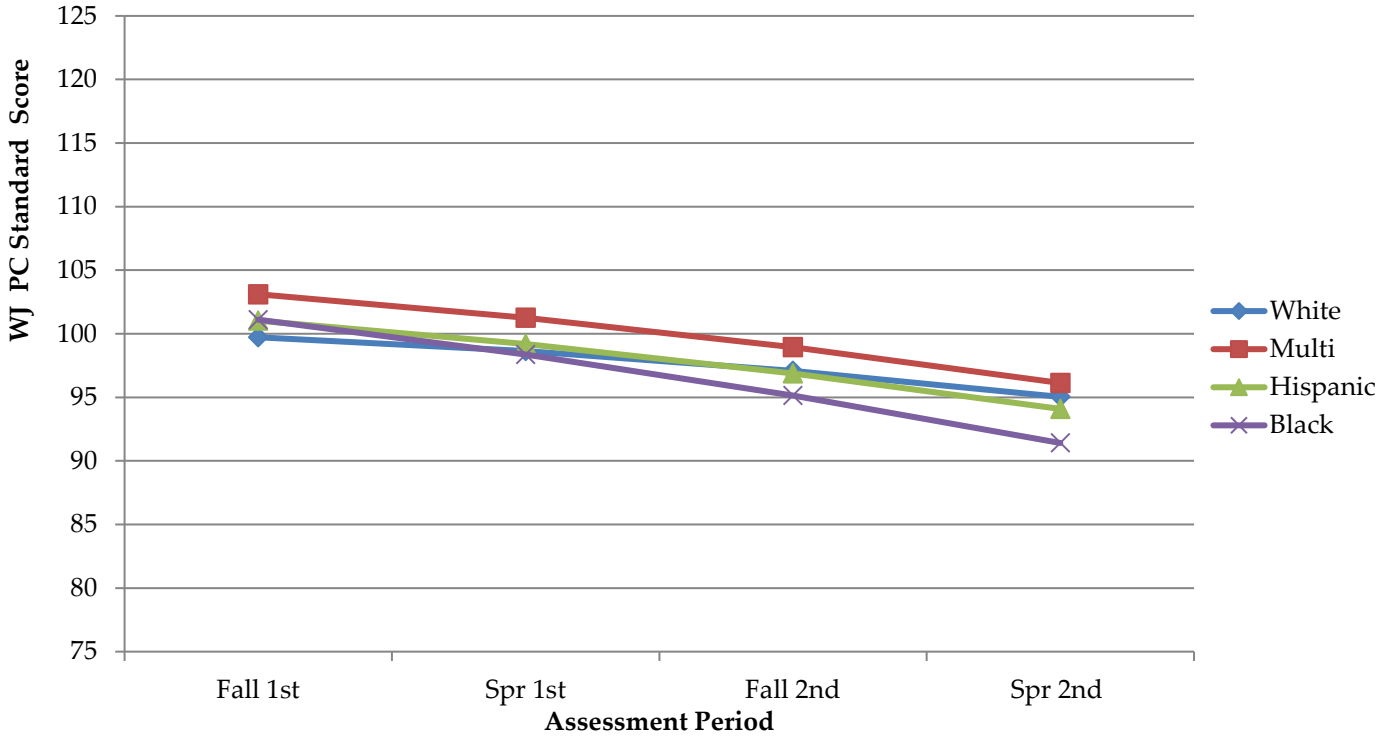


Figure 21. Growth in WJ-III Applied Problems by Race/Ethnicity

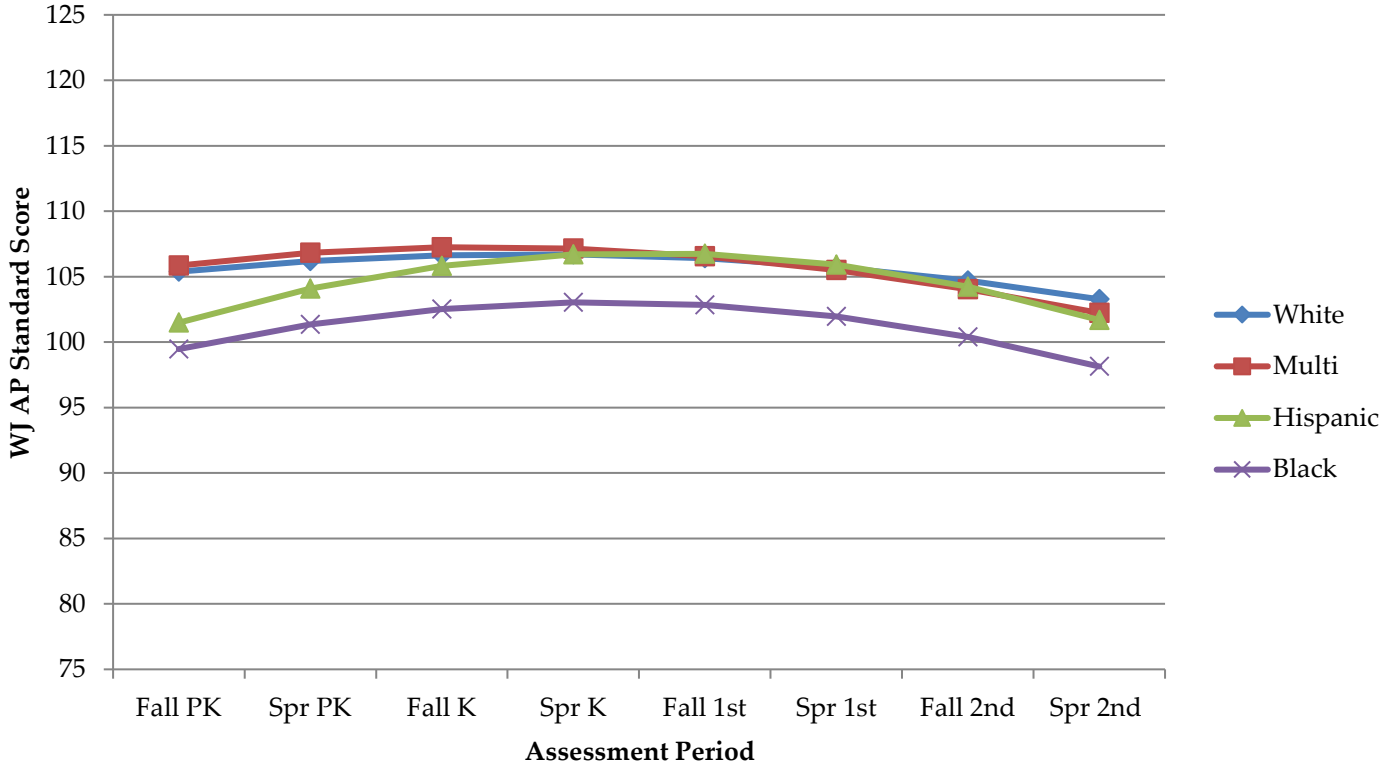


Figure 22. Growth in WJ-III Calculation by Race/Ethnicity

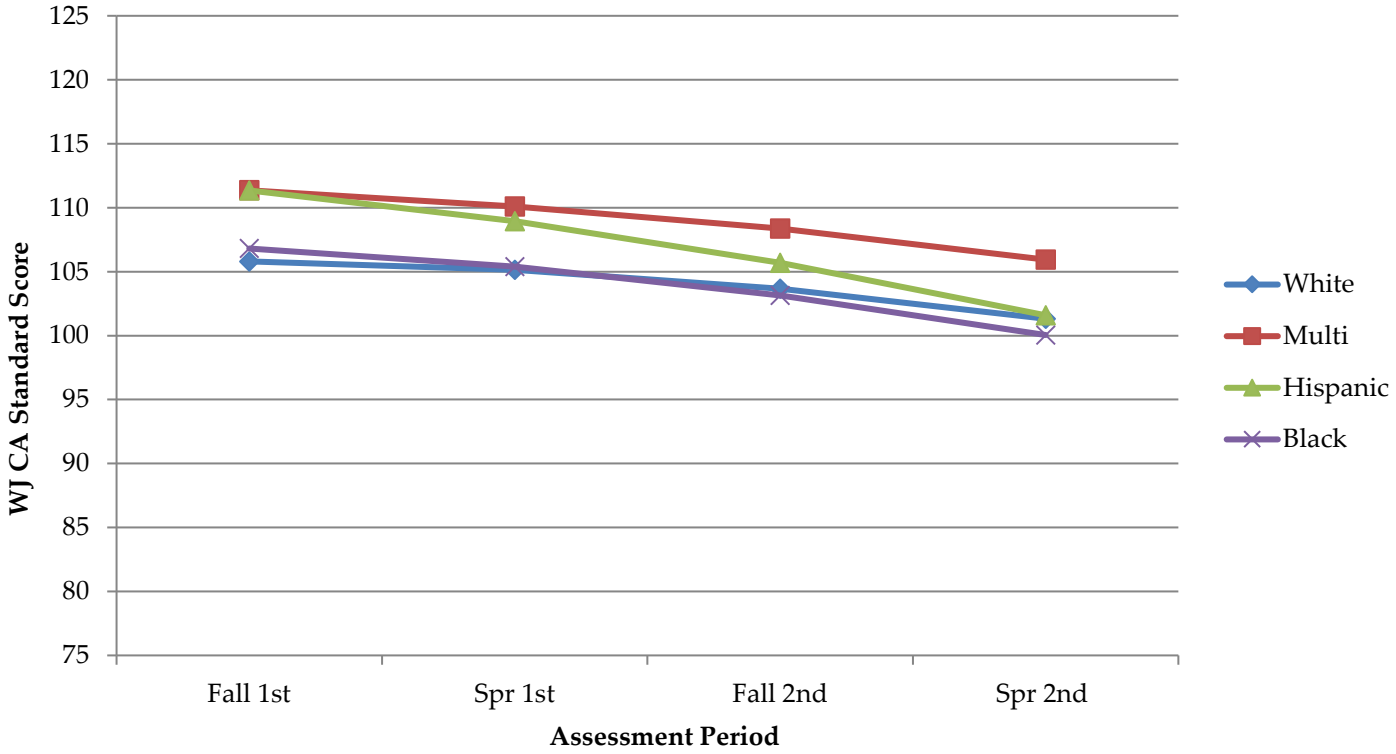


Figure 23. Growth in WJ-III Sound Awareness by Pre-K Program Type

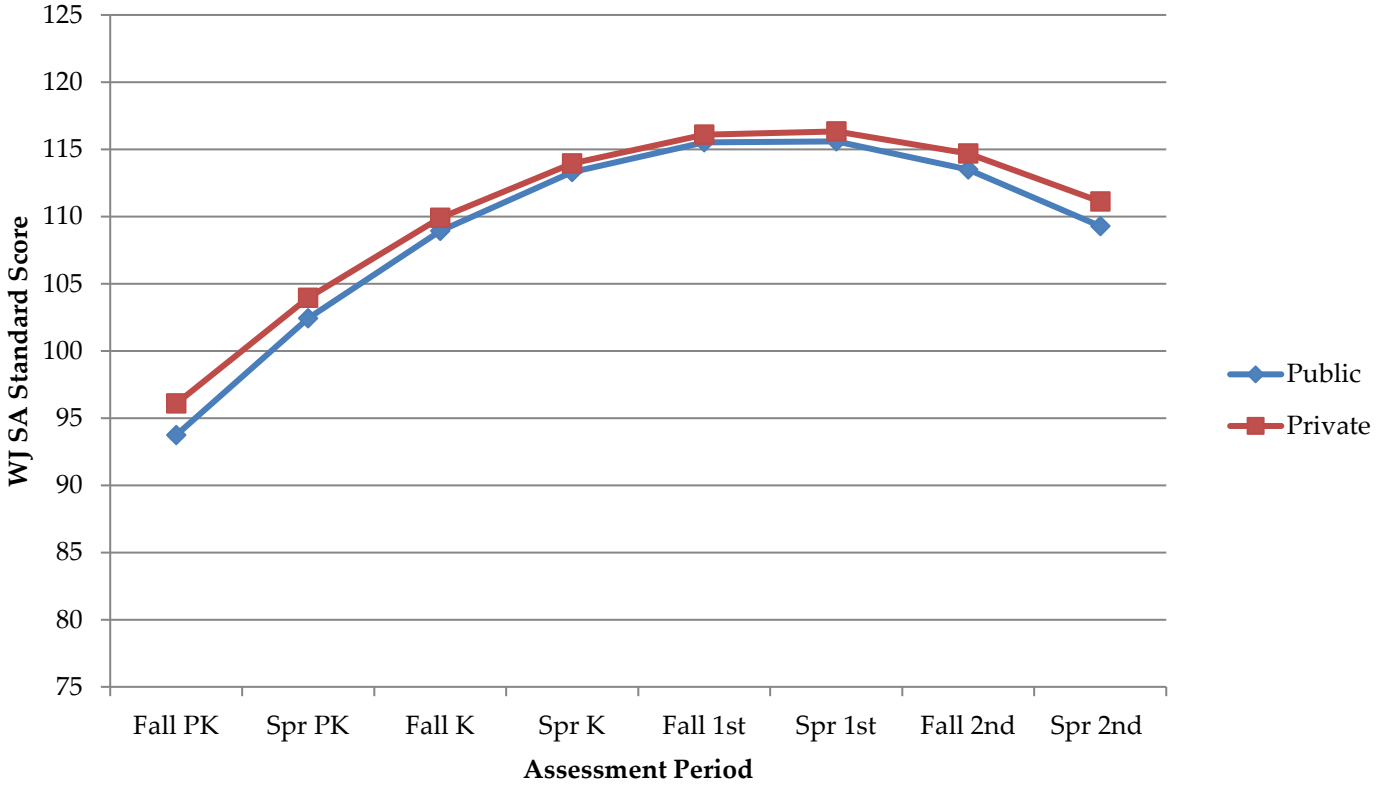


Figure 24. Growth in WJ-III Letter Word ID by Pre-K Program Type

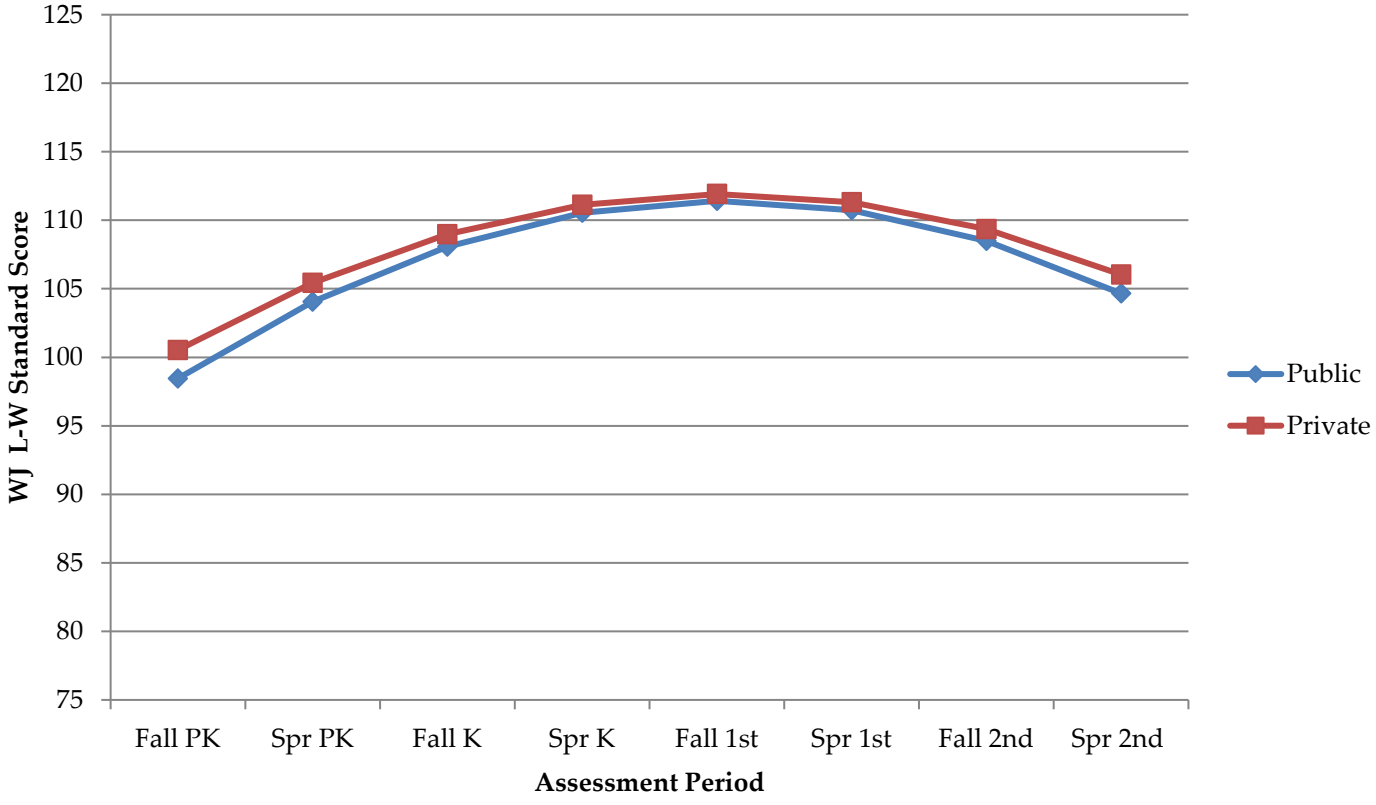


Figure 25. Growth in WJ-III Word Attack by Pre-K Program Type

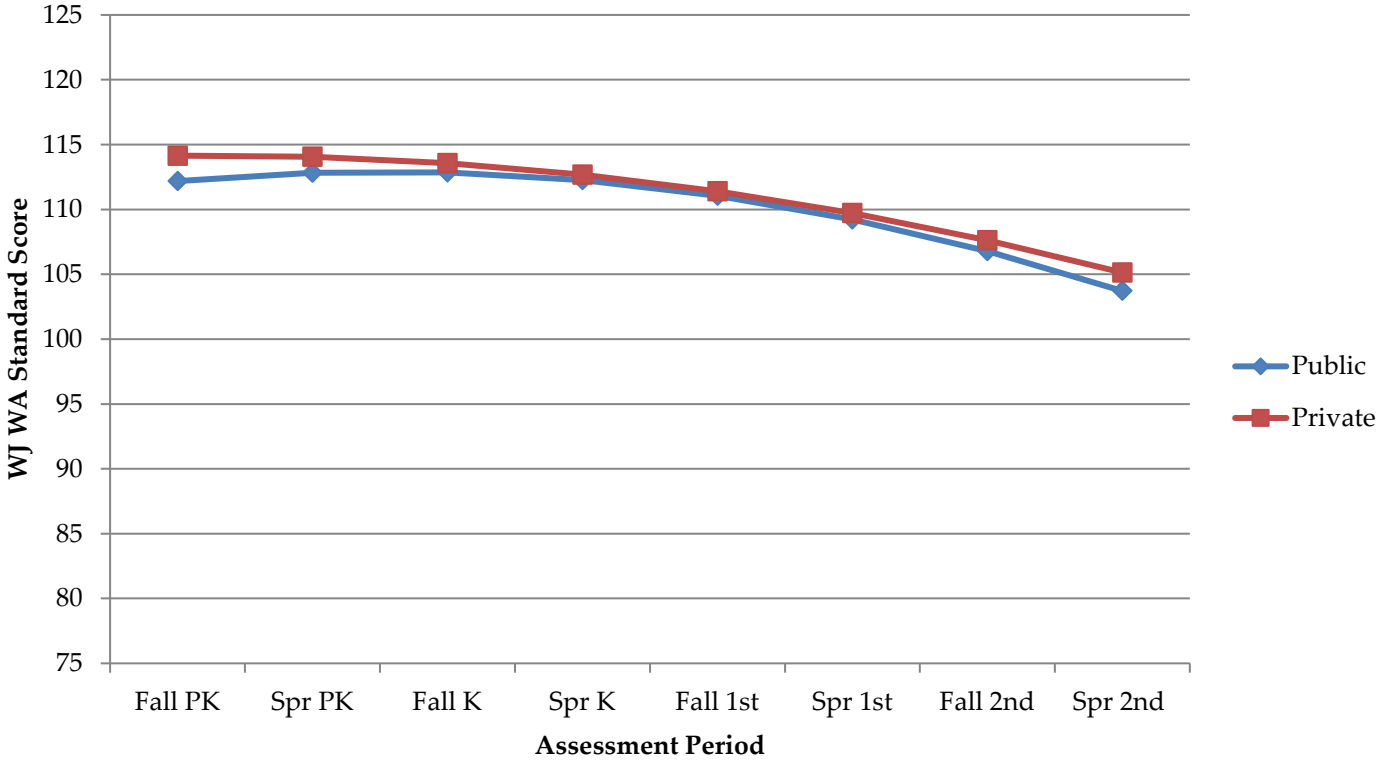


Figure 26. Growth in WJ-III Calculation by Pre-K Program Type

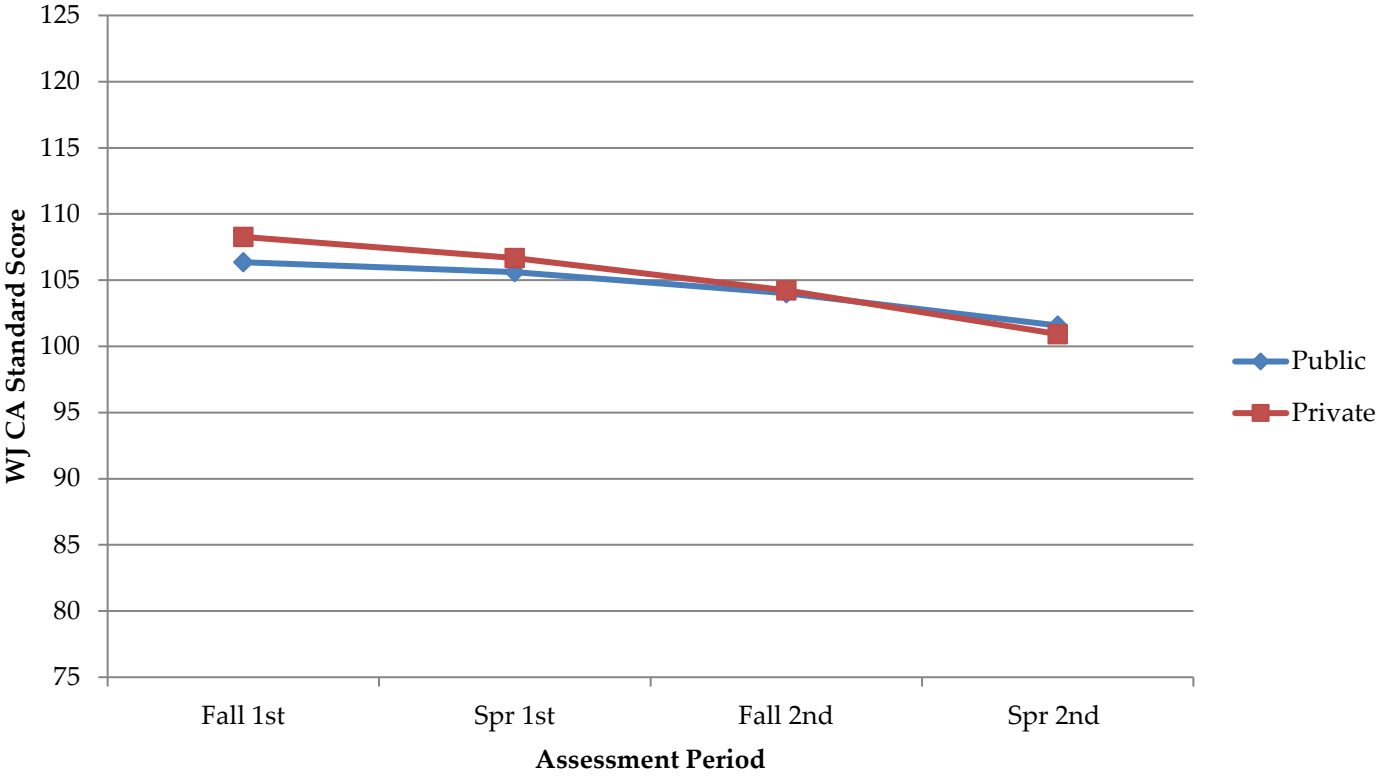


Figure 27. Growth in SSiS Problem Behavior by Pre-K Program Type

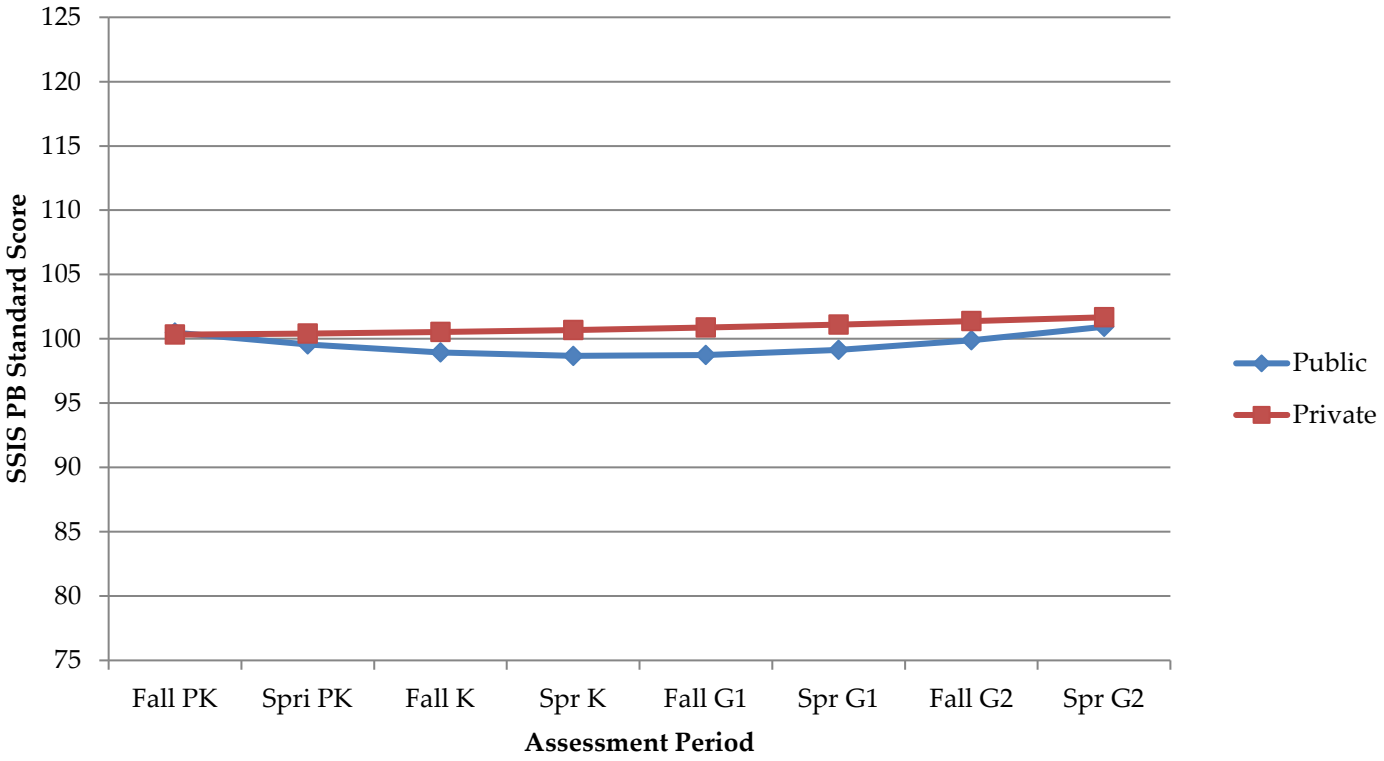


Figure 28. Growth in WJ-III Sound Awareness by K-2 CLASS

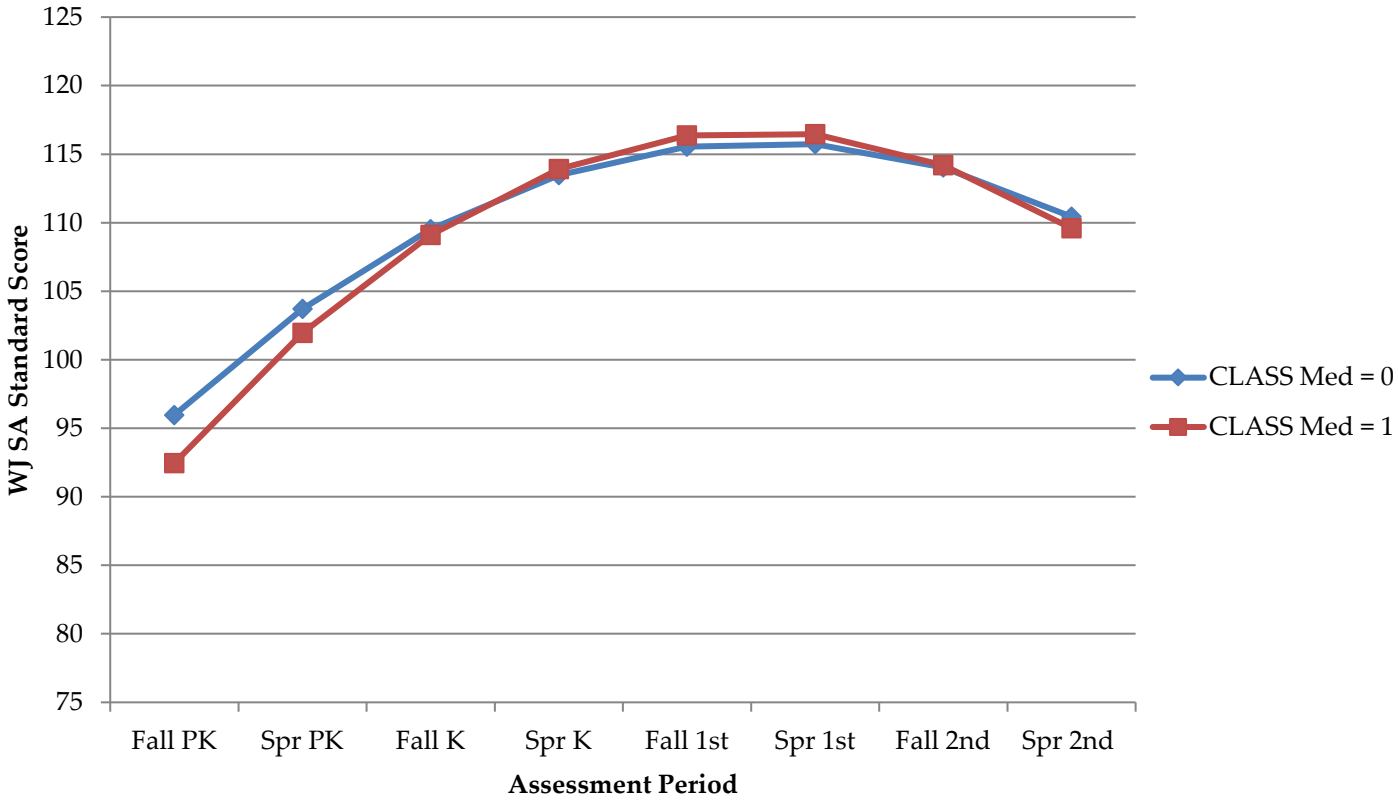


Figure 29. DLL Growth in WJ-III Picture Vocabulary by English Proficiency

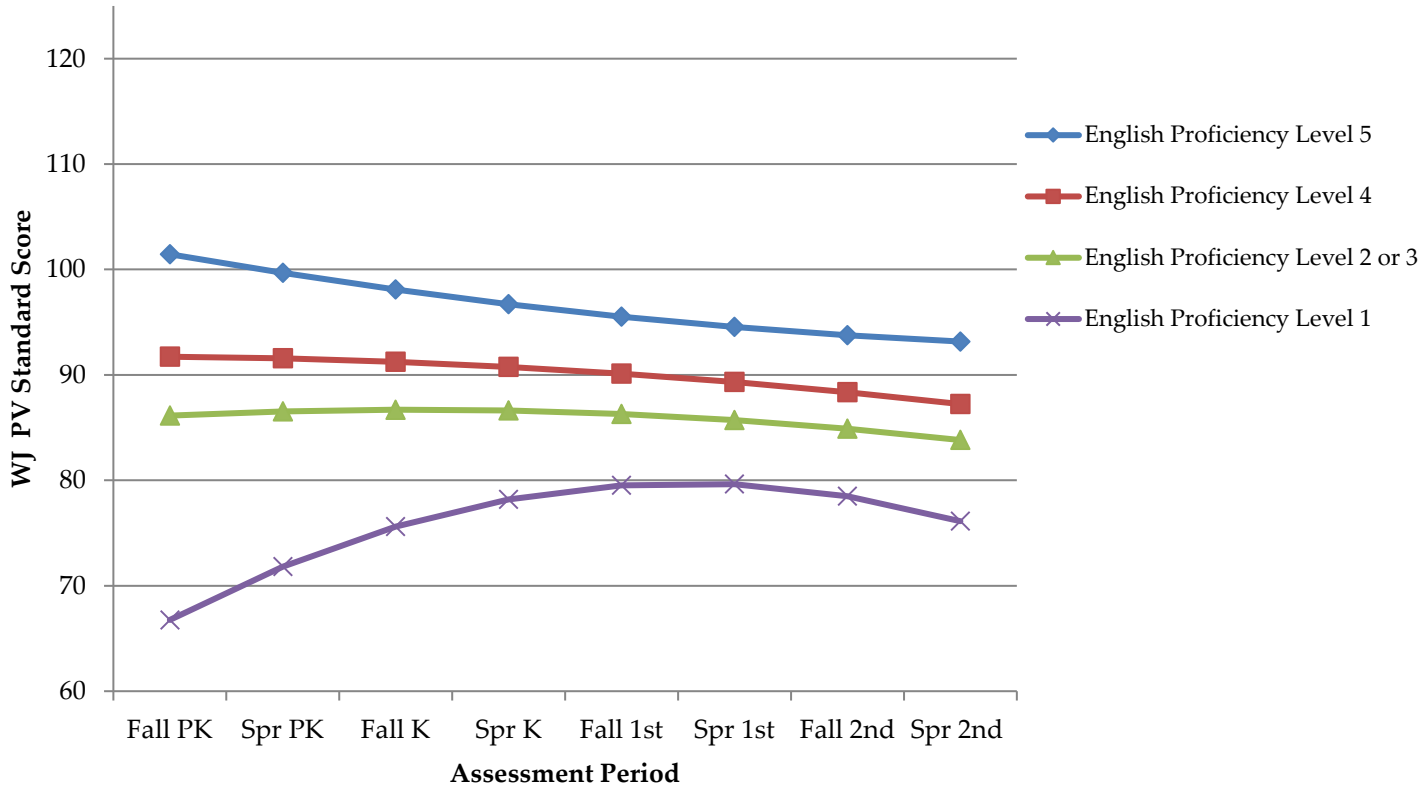


Figure 30. DLL Growth in WJ-III Sound Awareness by English Proficiency

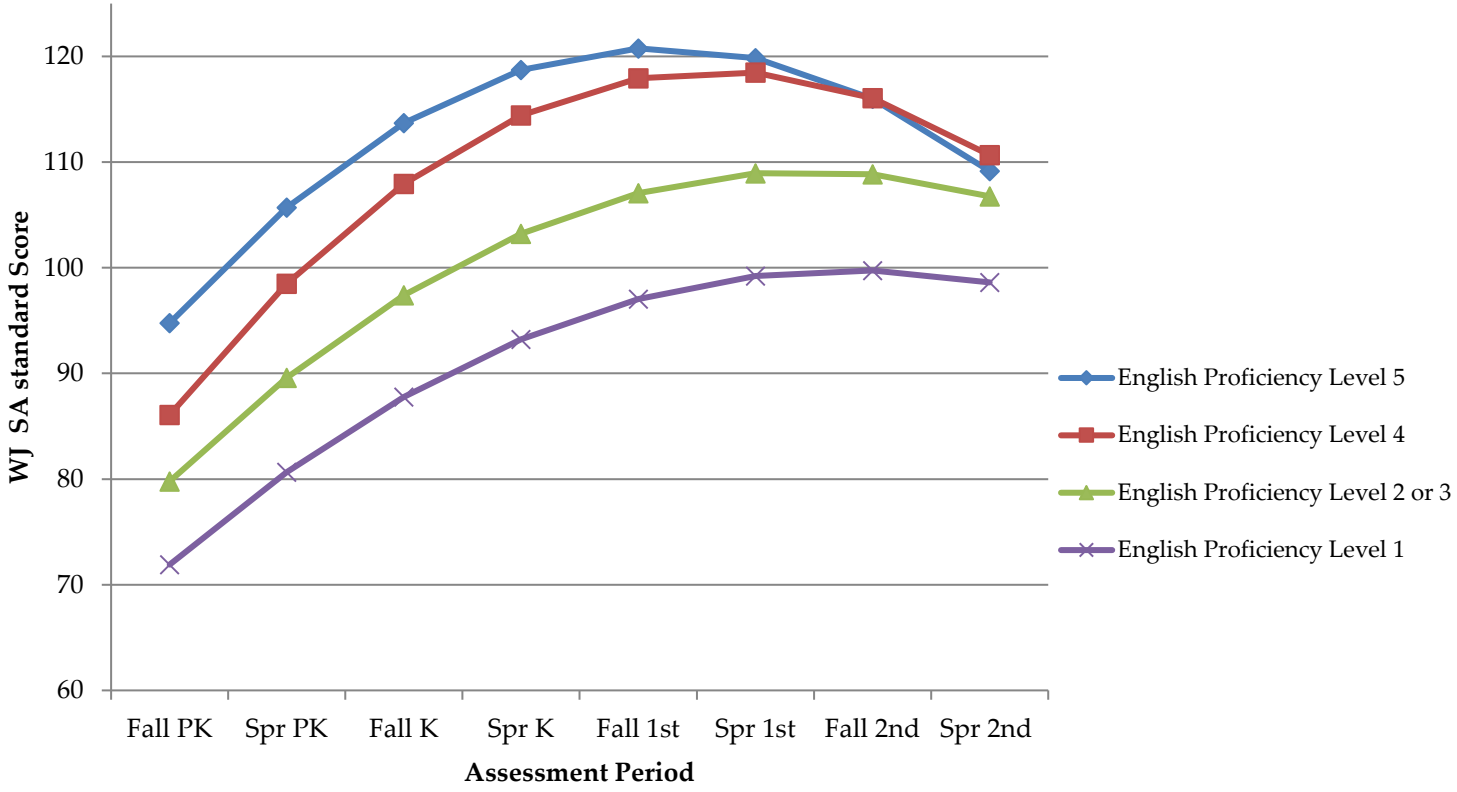


Figure 31. DLL Growth in WJ-III Applied Problems by English Proficiency

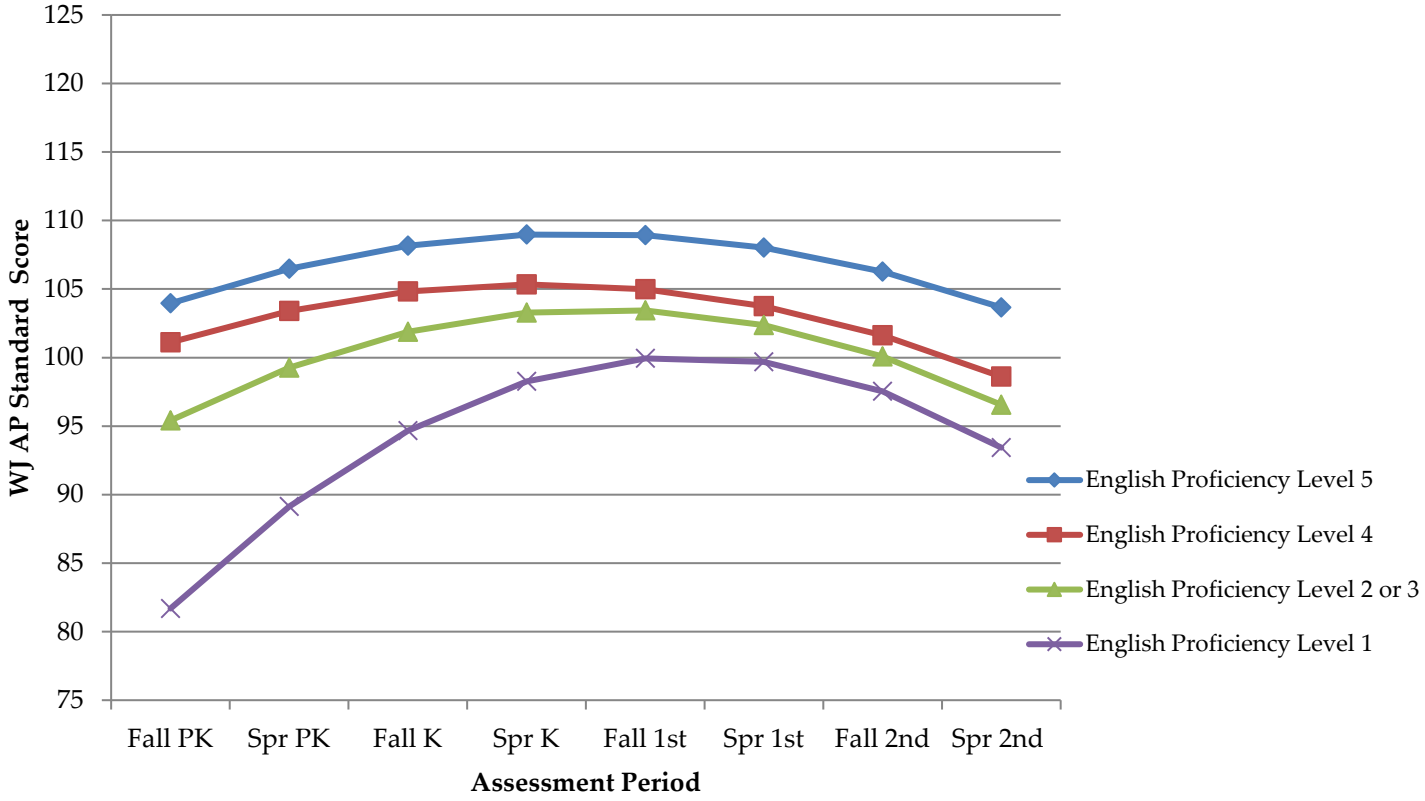


Figure 32. DLL Growth in WJ-III Letter Word Identification by Pre-K CLASS

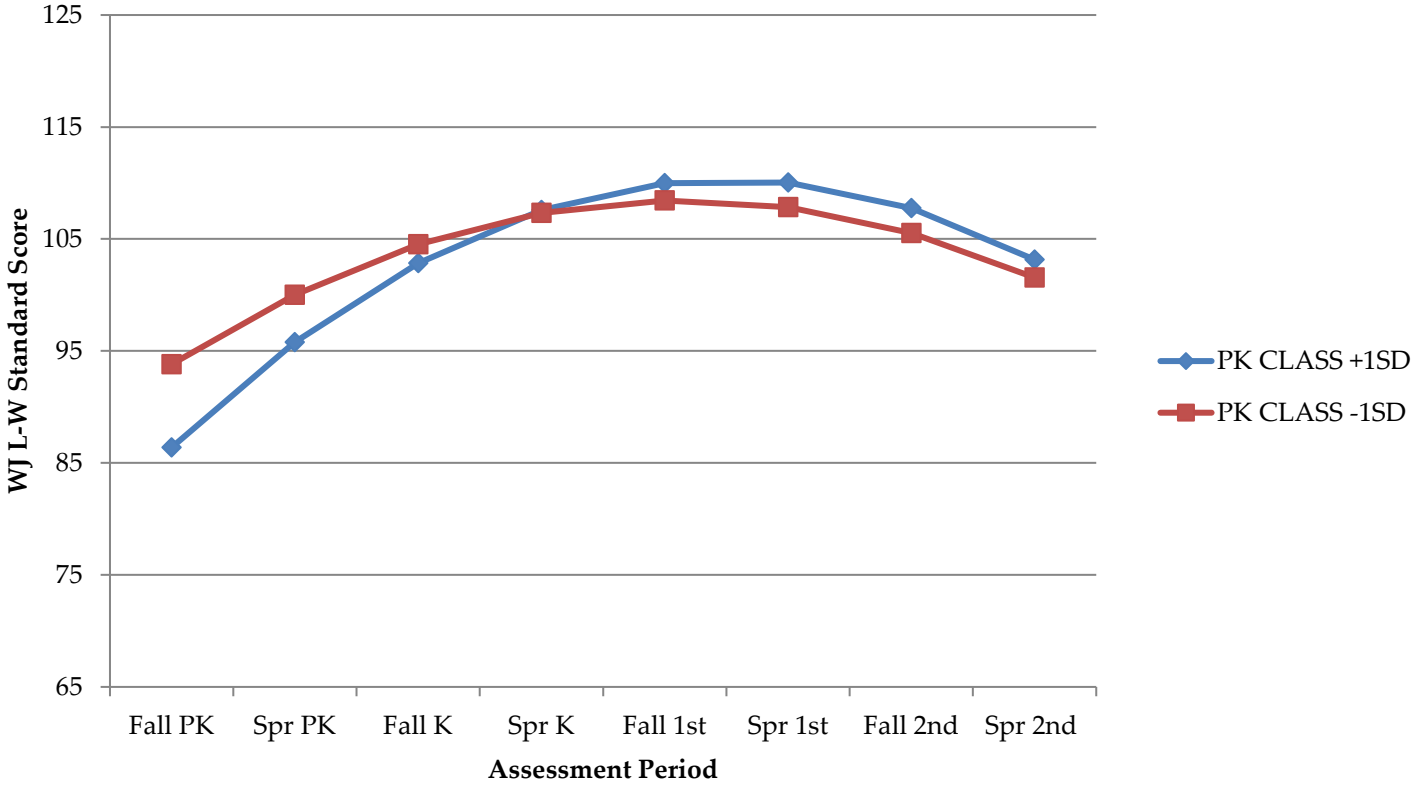


Figure 33. DLL Growth in WJ-III Picture Vocabulary by K-2 CLASS

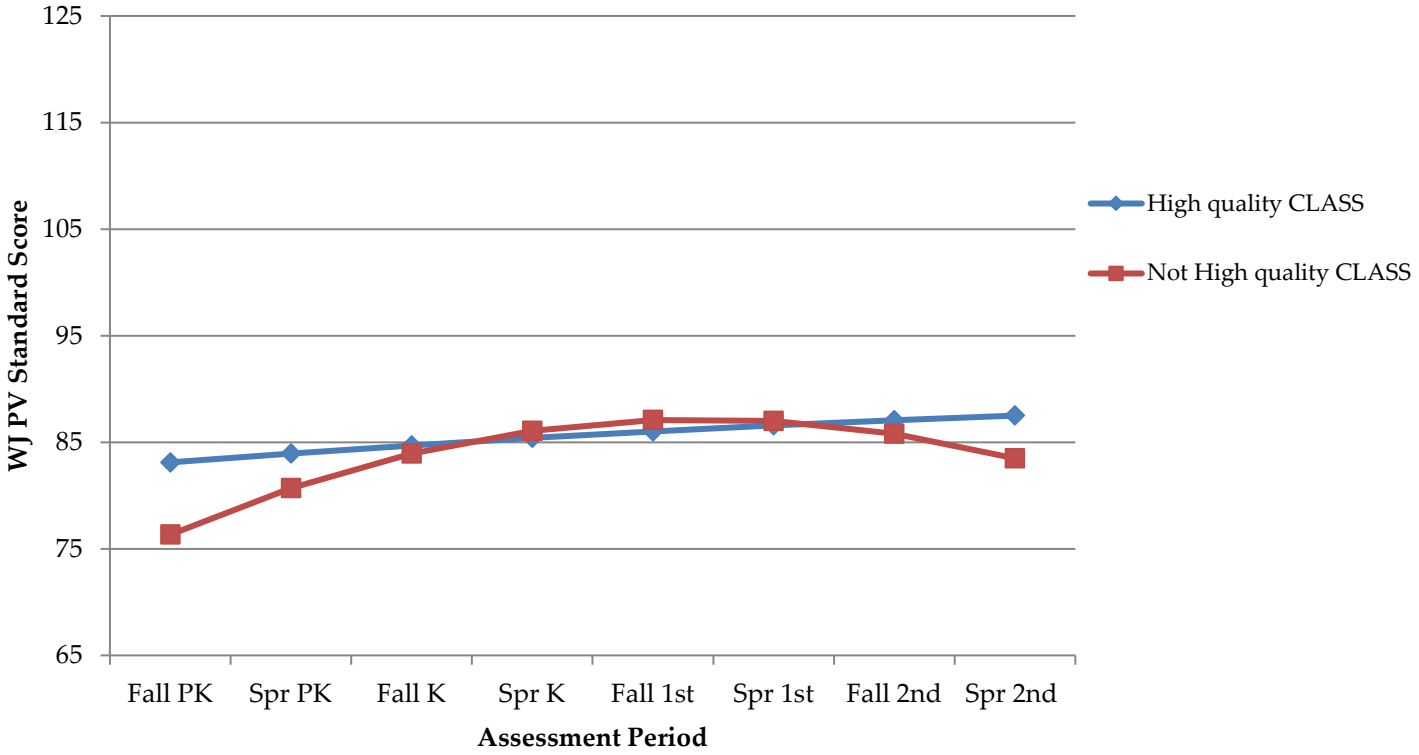


Figure 34. DLL Growth in WJ-III Calculation by K-2 CLASS

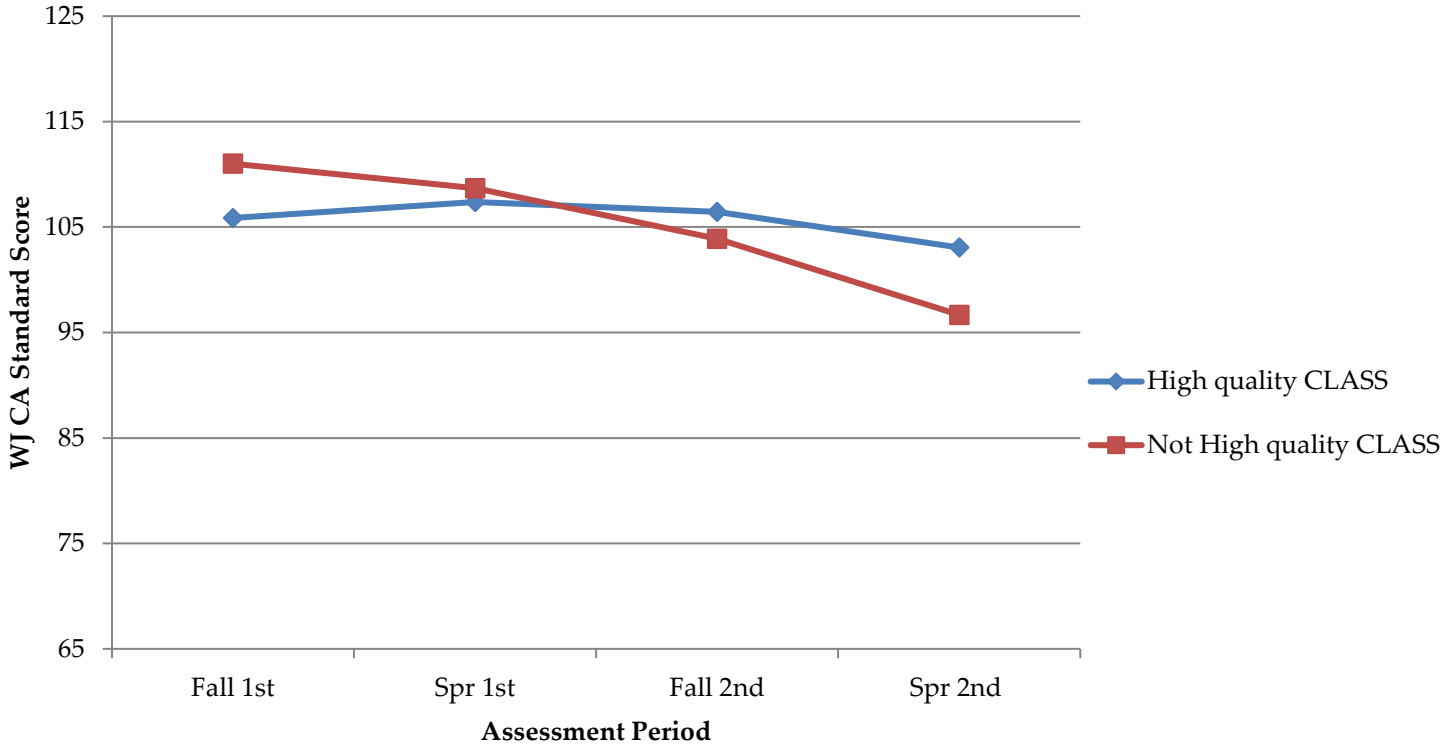


Figure 35. DLL Growth in WJ-III Picture Vocabulary by K-2 CLASS

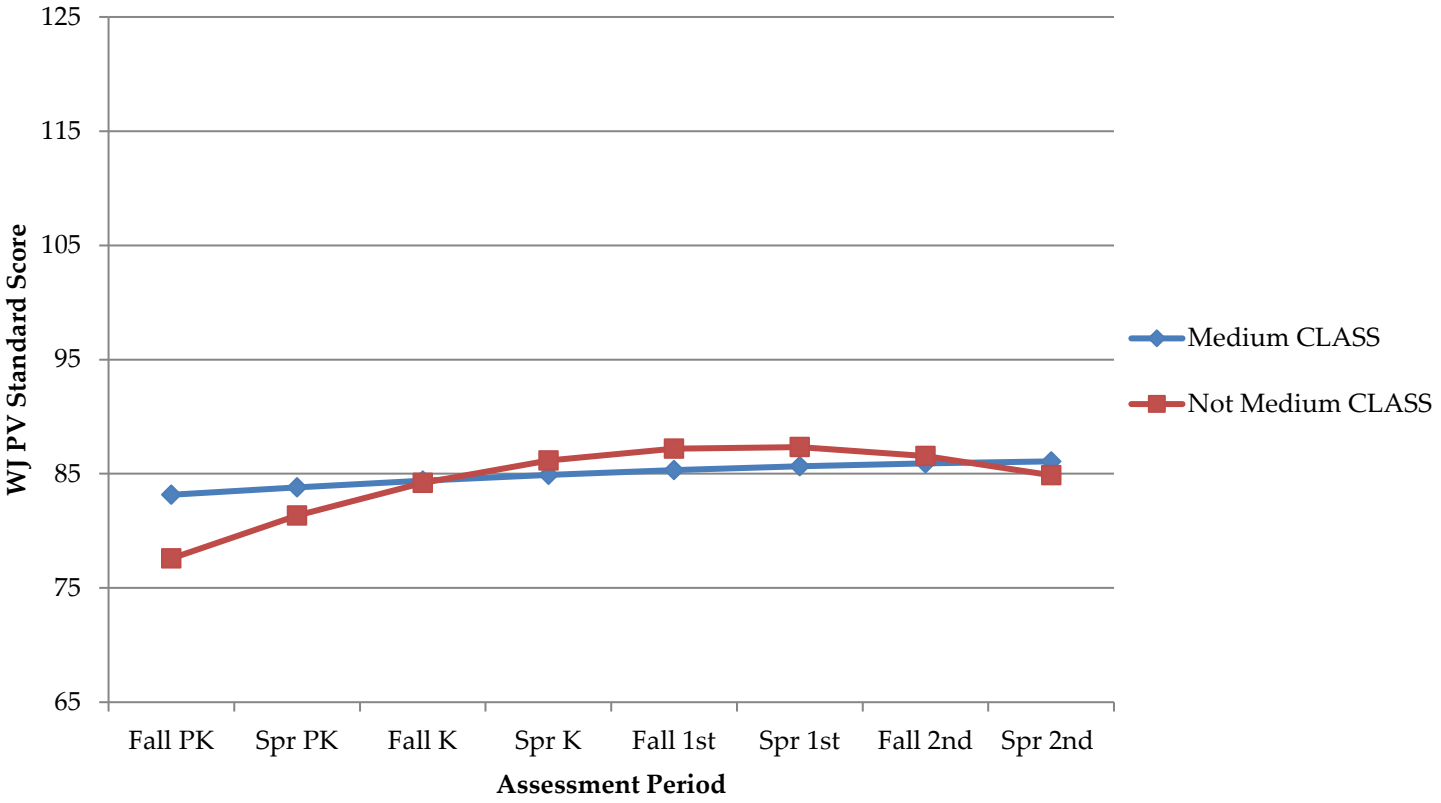


Figure 36. DLL Growth in WJ-III Passage Comprehension by K-2 CLASS

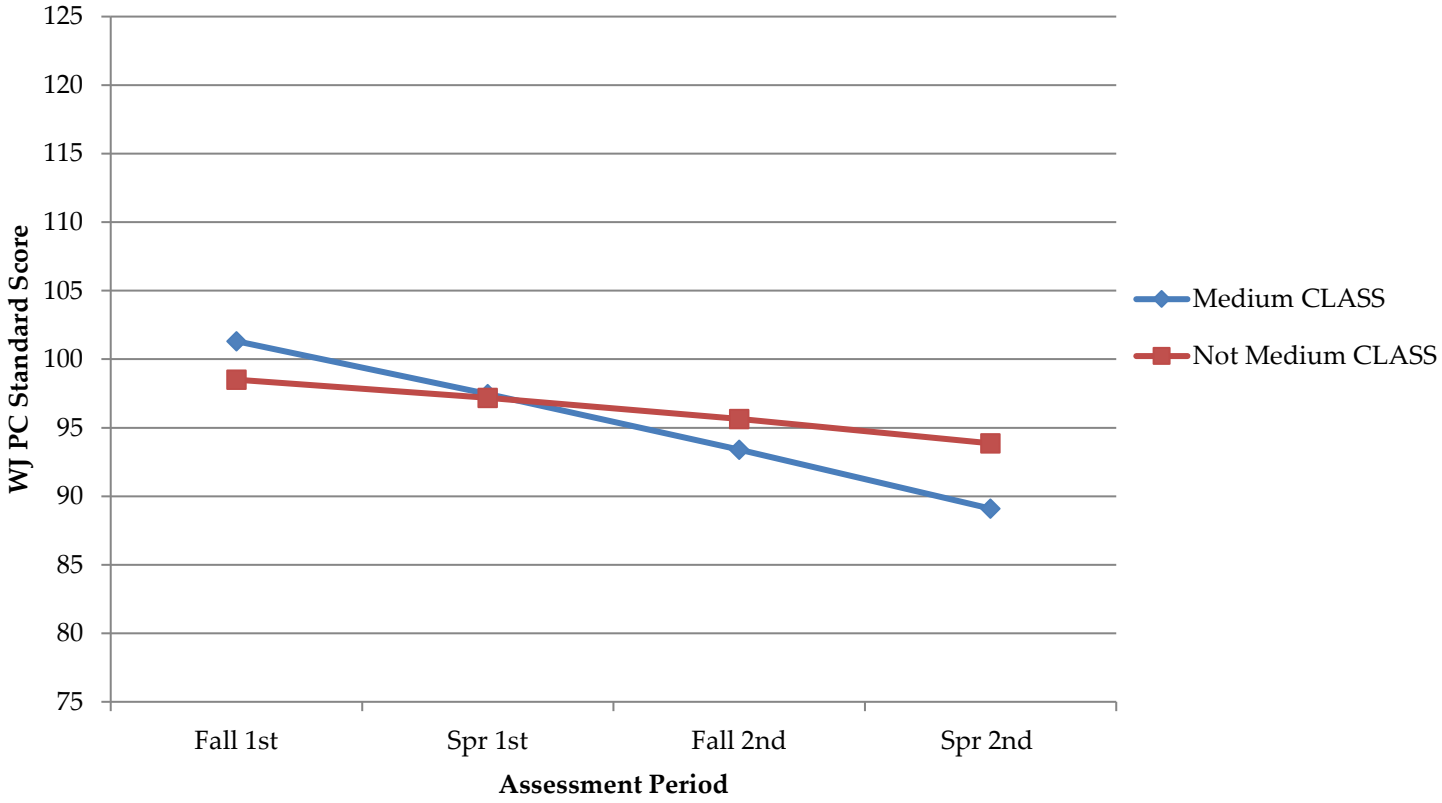


Figure 37. DLL Growth in Bat-III Sound Awareness by Spanish Proficiency

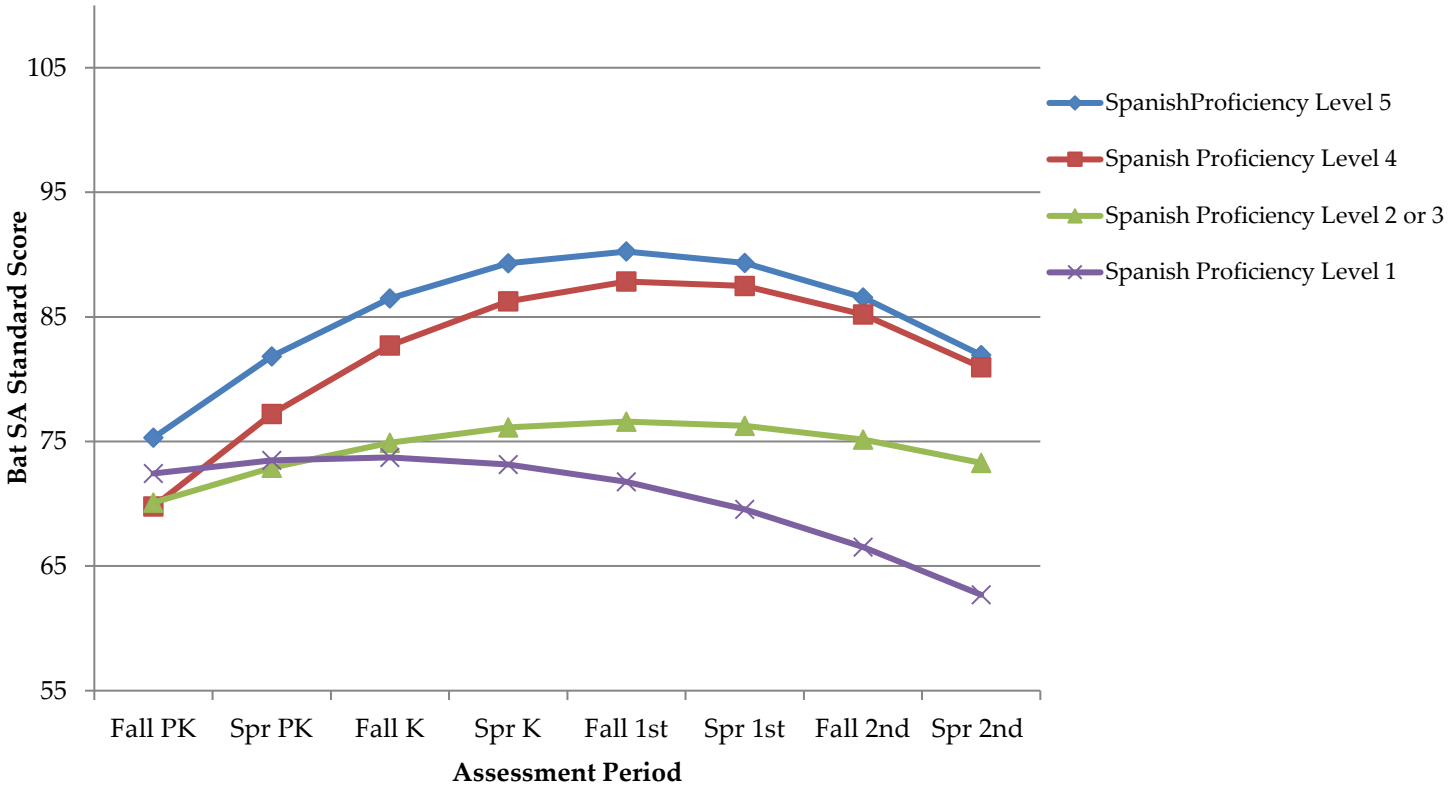


Figure 38. DLL Growth in Bat-III Letter Word ID by Spanish Proficiency

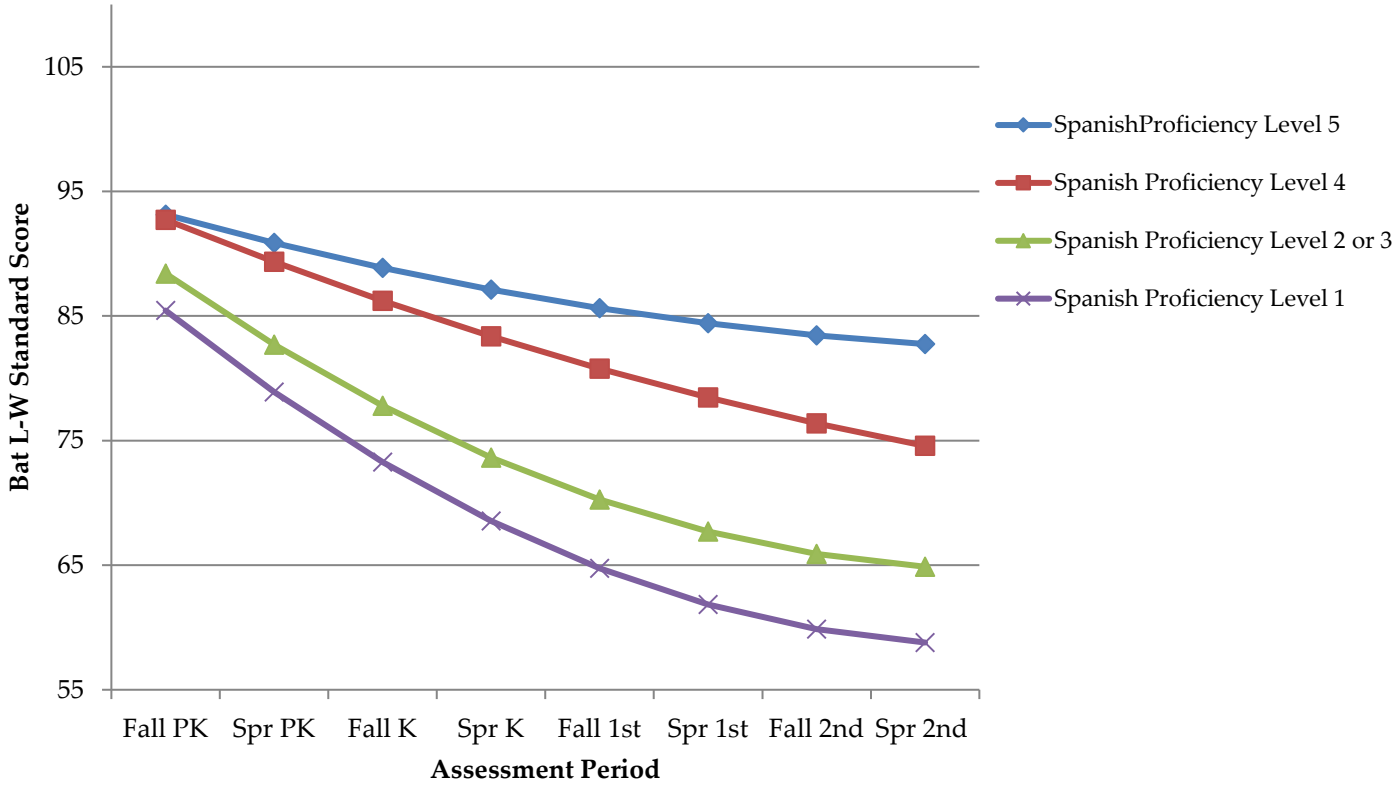


Figure 39. DLL Growth in Bat-III Word Attack by Spanish Proficiency

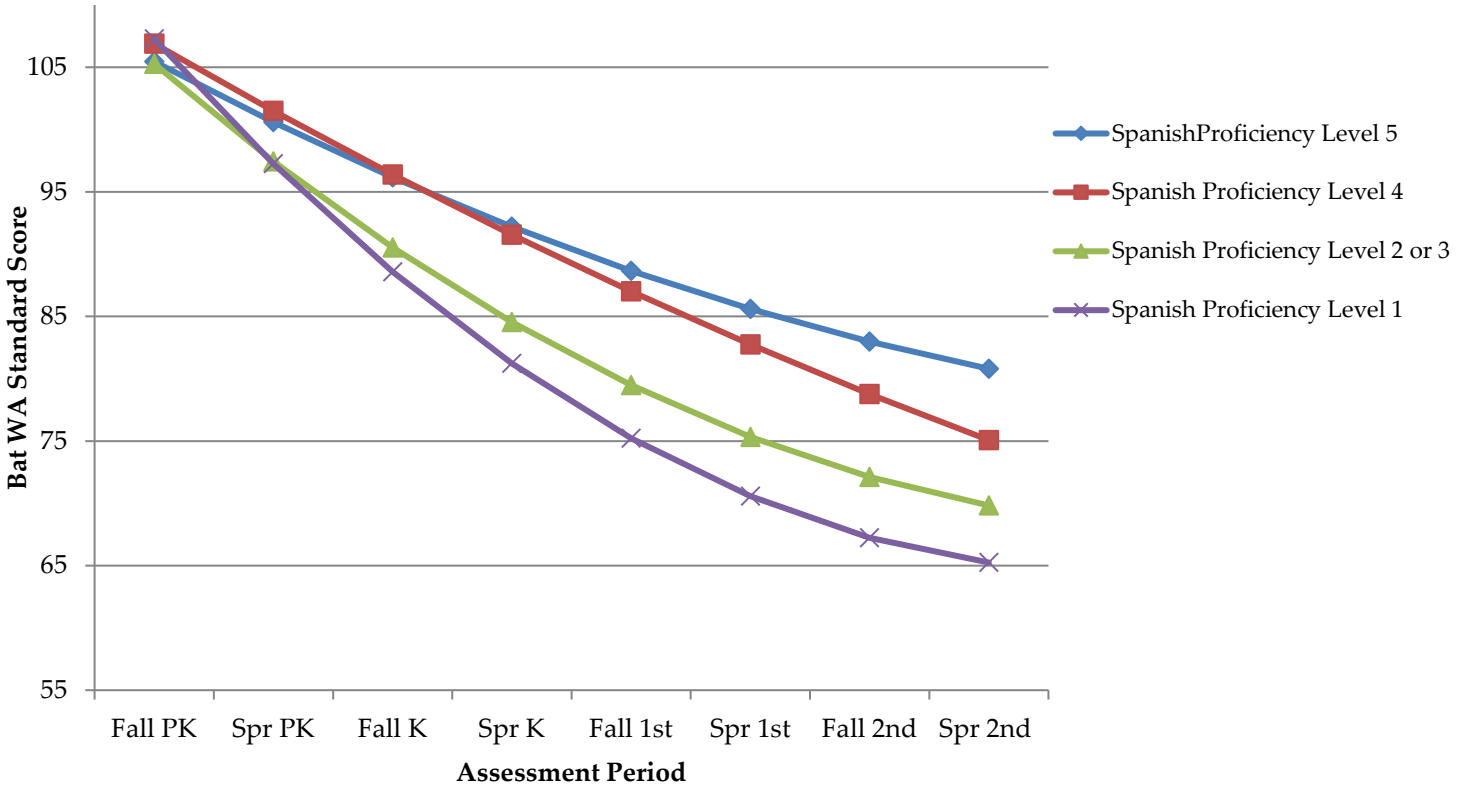


Figure 40. DLL Growth in Bat-III Applied Problems by Pre-K CLASS

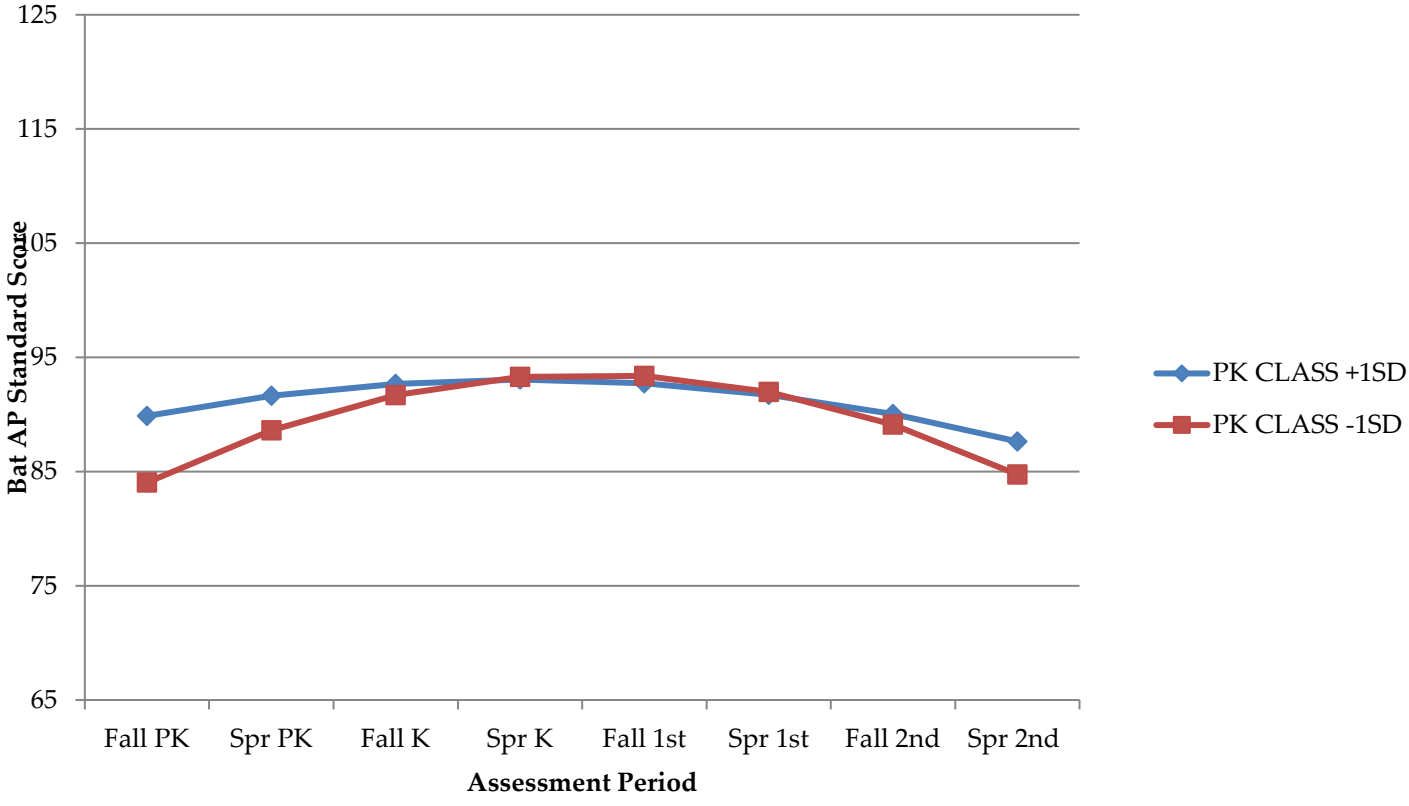


Figure 41. DLL Growth in Spanish Forward Digit Span by Pre-K CLASS

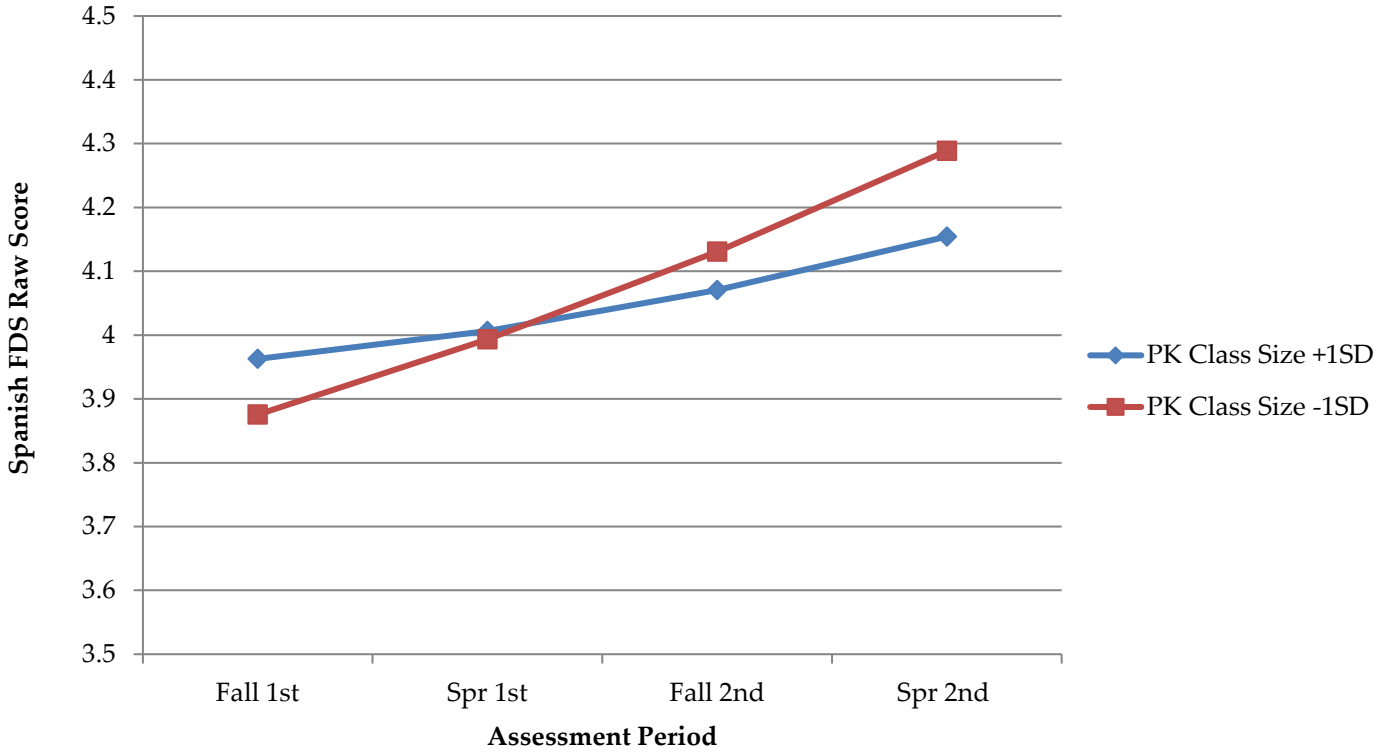


Table 18. Classroom Assessment Scoring System (CLASS) Scores (Pre-K – 2nd Grade)

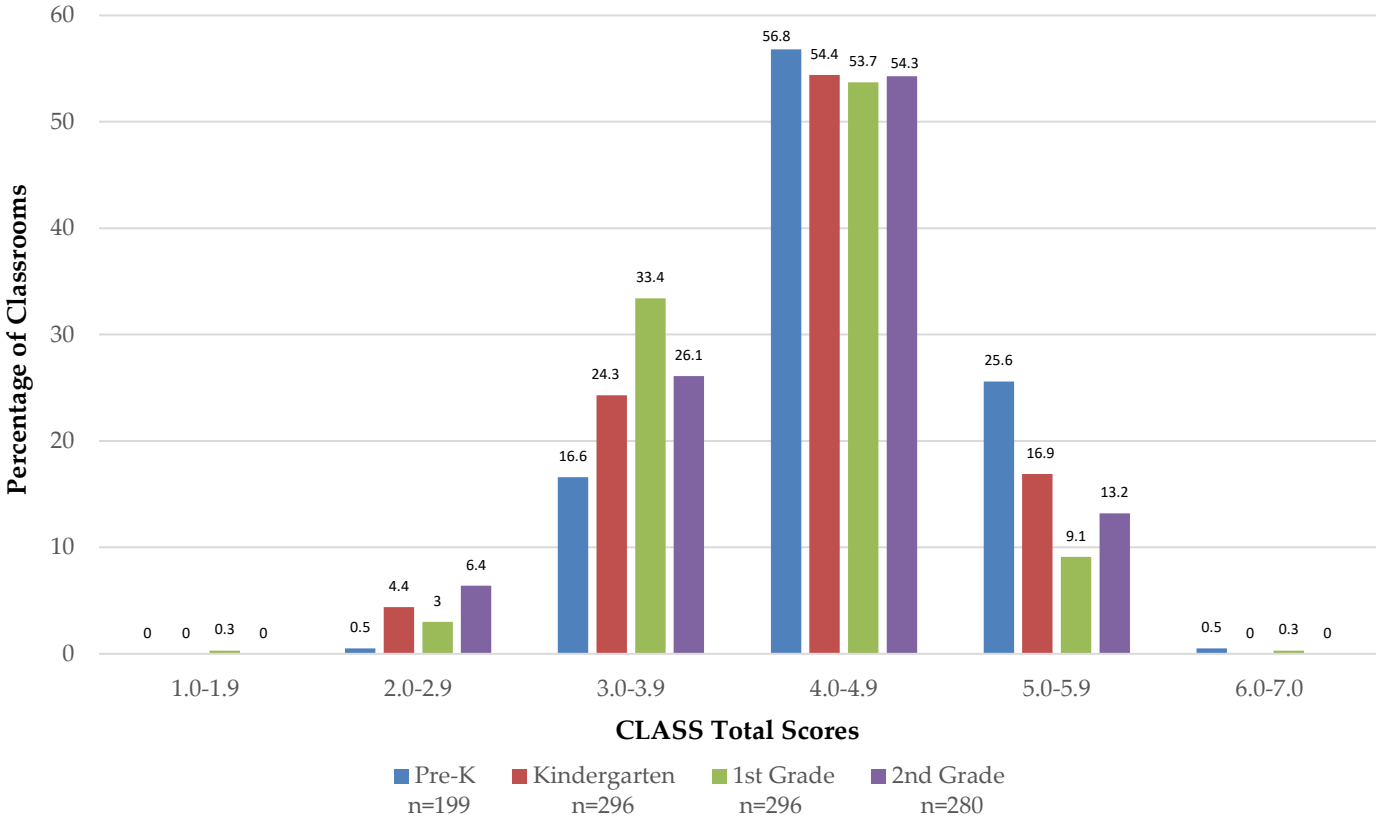
CLASS Item[a]	Pre-K n=199			Kindergarten n=296			1st Grade n=296			2nd Grade n=280			T-test Results
Domain/ Dimension	Mean	(SD)	Rangeb	Mean	(SD)	Rangea	Mean	(SD)	Rangea	Mean	(SD)	Rangea	Sig Group Comparisons
CLASS Total Score	4.6	(0.6)	3.0-6.5	4.3	(0.7)	2.2-5.9	4.2	(0.7)	1.7-6.3	4.2	(0.7)	2.5-5.9	PK>K, 1st, 2nd *** K>1st **
Emotional Support	5.7	(0.7)	3.6–6.9	5.2	(0.8)	1.7–6.9	5.1	(0.7)	2.7-6.9	5.3	(0.8)	2.9-6.9	PK>K, 1st, 2nd *** 2nd>1st **
Positive climate	5.9	(0.9)	3.0–7.0	5.3	(1.1)	1.8–7.0	5.1	(1.0)	1.8-7.0	5.4	(1.0)	2.4-7.0	
Negative climatec	1.2	(0.4)	1.0–3.4	1.4	(0.7)	1.0–5.8	1.3	(0.5)	1.0-4.8	1.3	(0.5)	1.0-3.4	
Teacher sensitivity	5.5	(1.0)	2.2–7.0	5.1	(1.1)	1.3–7.0	4.9	(1.1)	1.4-7.0	5.2	(1.1)	2.0-7.0	
Regard for student perspectives	4.7	(1.0)	1.6–6.8	3.7	(1.0)	1.5–6.4	3.8	(0.9)	1.0-6.4	3.9	(0.9)	1.0-6.8	
Classroom Organization	5.5	(0.8)	3.1–6.9	5.3	(0.8)	2.7–6.9	5.1	(0.8)	1.3-6.9	5.2	(0.9)	2.9-6.8	PK>K*, 1st, 2nd *** K>1st ***
Behavior management	5.8	(1.0)	2.8–7.0	5.6	(1.0)	1.6–7.0	5.4	(1.0)	1.2-7.0	5.5	(1.1)	2.2-7.0	
Productivity	5.9	(0.7)	3.4–7.0	5.6	(0.9)	2.6–7.0	5.3	(0.9)	1.4-7.0	5.5	(0.9)	3.0-7.0	
Instructional learning formats	4.7	(0.9)	2.0–6.8	4.8	(1.0)	2.0–7.0	4.5	(0.9)	1.2-6.8	4.5	(1.0)	1.2-6.8	
Instructional Support	2.6	(0.7)	1.1–5.9	2.5	(0.8)	1.0–4.9	2.4	(0.7)	1.0-5.2	2.2	(0.7)	1.0-4.6	PK>1st *, 2nd *** K>2nd *** 1st>2nd *
Concept development	2.5	(0.8)	1.0–5.6	2.4	(0.8)	1.0–4.6	2.2	(0.7)	1.0-5.4	2.1	(0.7)	1.0-4.4	
Quality of feedback	2.6	(0.9)	1.0–6.0	2.6	(1.0)	1.0–5.8	2.5	(0.9)	1.0-5.4	2.3	(0.8)	1.0-5.2	
Language modeling	2.6	(0.8)	1.0–6.2	2.5	(0.9)	1.0–5.4	2.4	(0.8)	1.0-5.2	2.2	(0.8)	1.0-4.8	

[a] CLASS was used in Pre-K; CLASS K-3 was used in K-2nd grades.

[b] Total and Domain scores could range from 1.0–7.0; Dimension scores could range from 1–7.

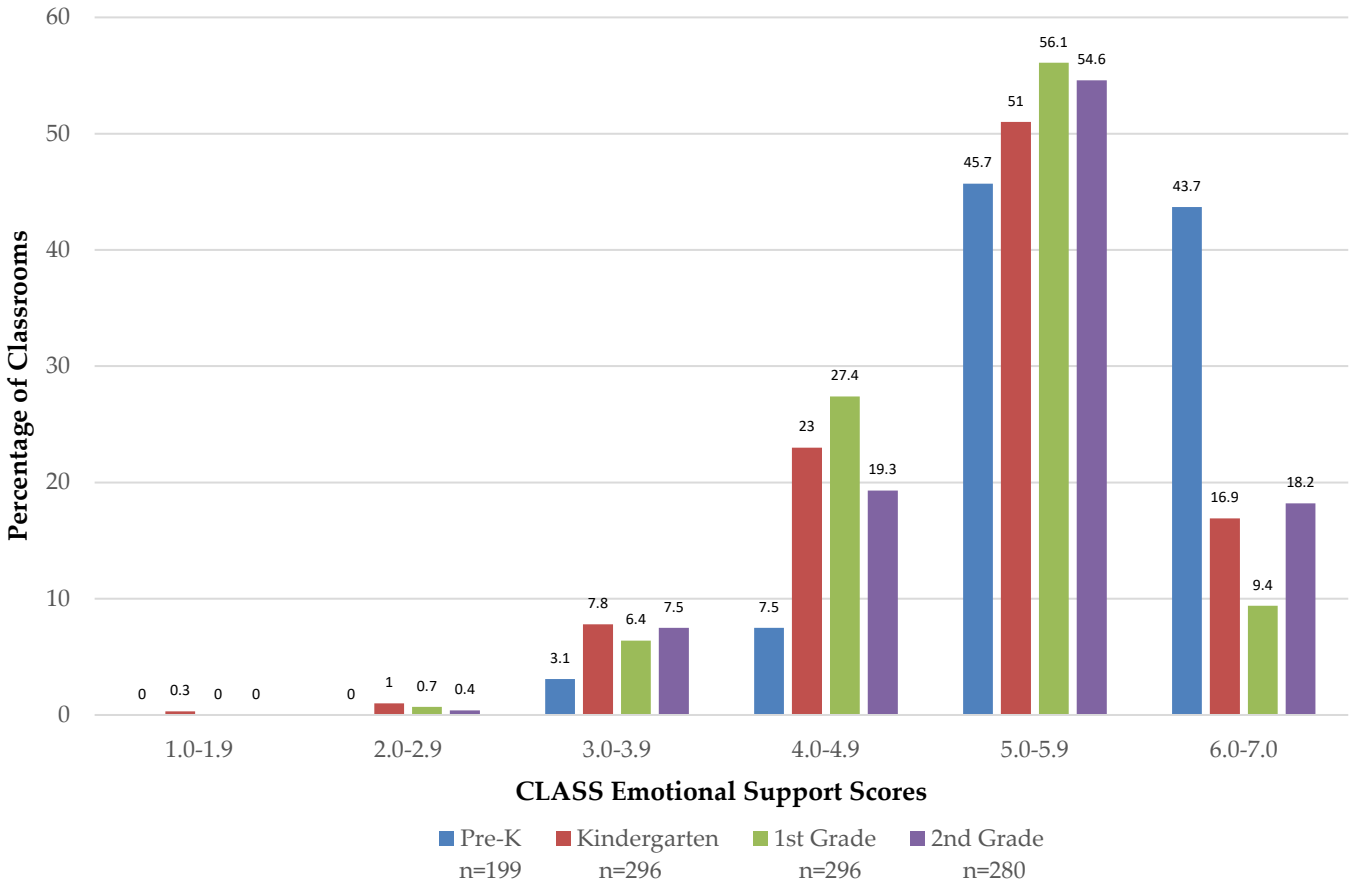
[c] Scoring is reversed for the Negative climate dimension before it is included in the calculation of the Emotional Support domain score.

Figure 42. CLASS Total Scores in Pre-K – 2nd Grade



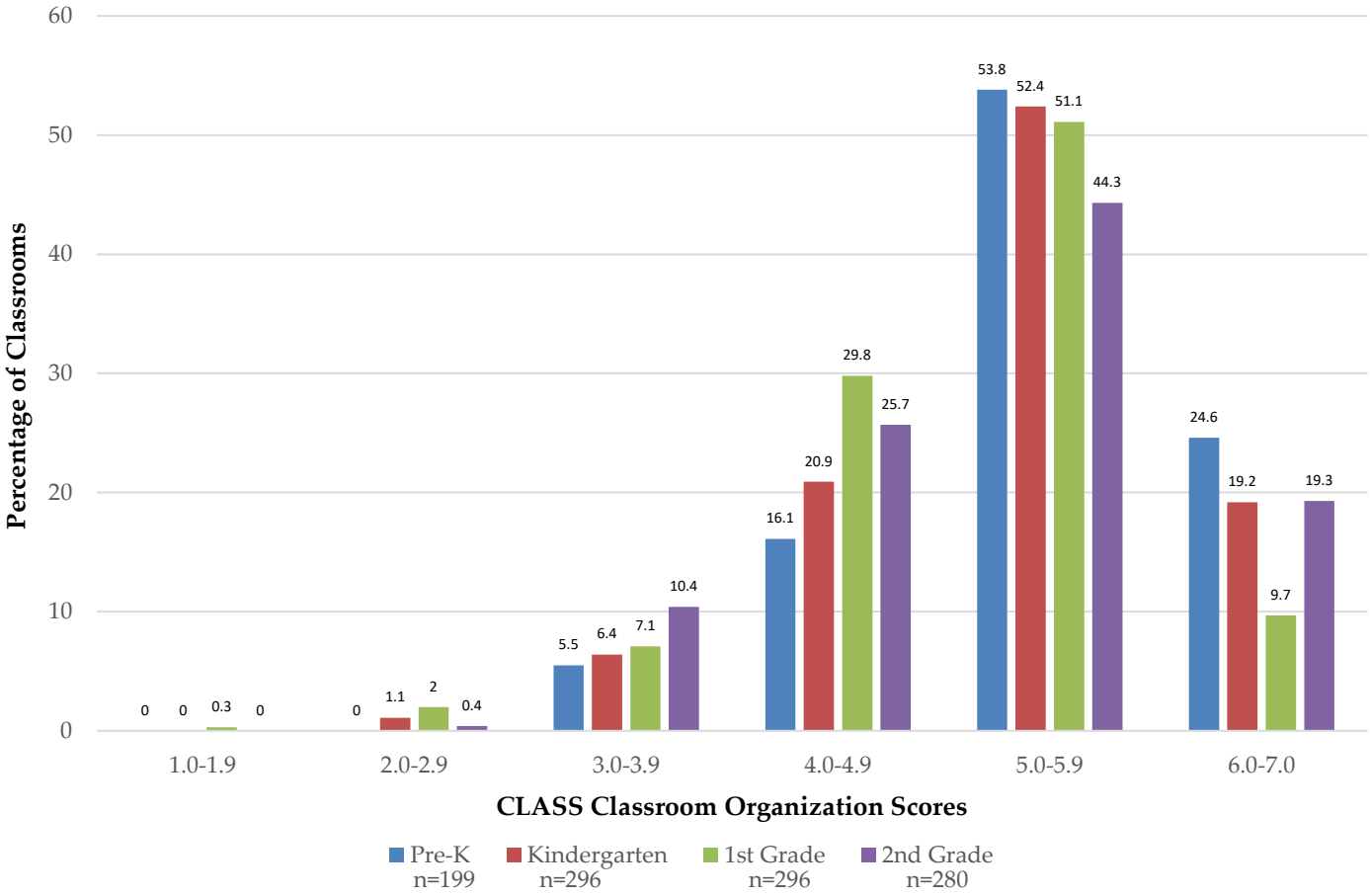
Note: CLASS was used in Pre-K; CLASS K-3 was used in K-2nd grades.

Figure 43. CLASS Emotional Support Scores in Pre-K – 2nd Grade



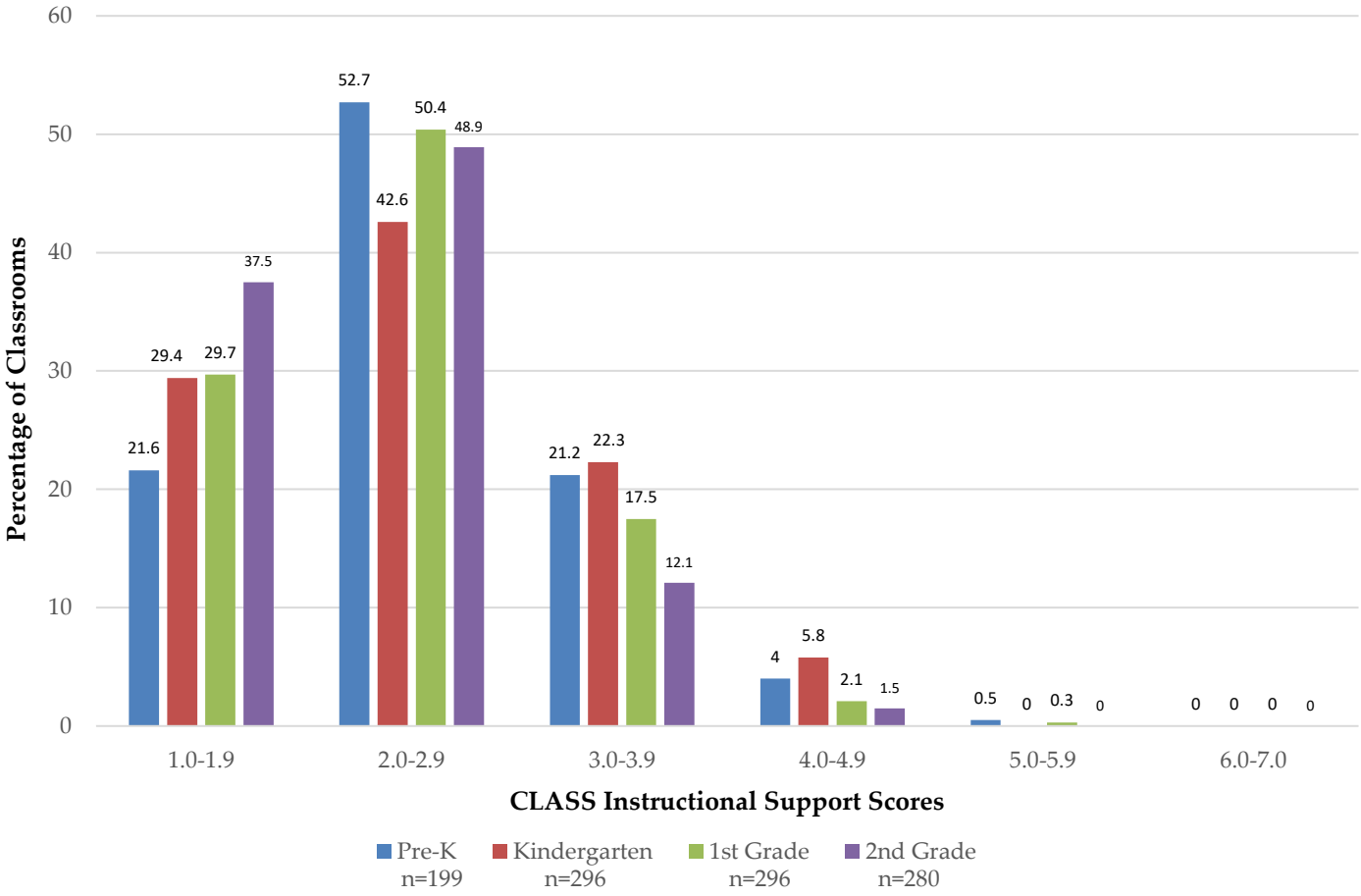
Note: CLASS was used in Pre-K; CLASS K-3 was used in K-2nd grades.

Figure 44. CLASS Classroom Organization Scores in Pre-K – 2nd Grade



Note: CLASS was used in Pre-K; CLASS K-3 was used in K-2nd grades.

Figure 45. CLASS Instructional Support Scores in Pre-K – 2nd Grade



Note: CLASS was used in Pre-K; CLASS K-3 was used in K-2nd grades.

Appendix: Child Outcomes Analysis Approach

Each of the outcome scores, collected in fall and spring each year from pre-k through second grade (or during first and second grade only for some outcomes), served as the dependent variables. A 3-level HLM was fit that estimated individual quadratic growth curves at level 1, linear effects for repeated child outcomes at level 2, and the pre-k classroom at level 3. Separate intercepts and slopes were estimated for each child at level 1 to describe individual differences in development over time and for each pre-k classroom at level 2 to account for the impact of pre-k experiences on individual development. Analyses tested whether developmental patterns varied by the child's race/ethnicity, language proficiency at entry to pre-k, IEP status, and family income (whether considered poor); pre-k class size and pre-k provider type (public vs private); and observed classroom quality in pre-k (total CLASS score) and in subsequent grades (high, medium, or low based on the CLASS). Each of these child, pre-k characteristics, and classroom quality variables was included as a main effect and crossed with linear change in all analyses and with quadratic change in analyses of outcomes collected from pre-k to second grade. Each coefficient describes the association between that term and the child outcome after adjusting for the other variables in the model. Thus, the interaction between a predictor and time-squared describes the degree to which quadratic change in that outcome differed for different levels of the predictor, adjusting for the associations between the other predictors and quadratic change. Similarly, the interaction between a predictor and time describes differences in linear rates of change during the pre-k year, adjusting for the other predictors (note linear change in subsequent years is computed from both the linear and quadratic coefficients).

The first set of models tested longitudinal growth over the four-year period for outcomes that were assessed from pre-k through second grade (WJ-III Picture Vocabulary, WJ-III Sound Awareness, WJ-III Letter-Word Identification, WJ-III Word Attack, WJ-III Applied Problems, SSiS Social Skills, and SSiS Problem Behaviors standard scores), and over the two-year period for outcomes that were added to the assessment battery in first grade (WJ-III Passage Comprehension and WJ-III Calculation standard scores, Forward Digit Span and Backward Digit Span raw scores). A separate model was estimated for each outcome.

Children were nested within pre-k classrooms, and repeated measures outcomes nested within child. Time (coded as 0, 1, 2, 3, 4, 5, 6, 7 for the eight time points from fall pre-k to spring second grade for longitudinal outcomes and 0, 1, 2, 3 for first- and second-grade only outcomes) represented the linear effect of growth on each outcome over time and Time-squared represented the quadratic effect of change in growth rates over time for outcomes assessed from pre-k through second grade. Both linear and quadratic growth effects were estimated, to allow for varying rates of change for each outcome. Predictors were included as main effects and in interactions with time – linear and quadratic for the first set of outcomes and linear for the second set of outcomes.

The base model included three sets of covariates:

1) Child/family characteristics – gender (F=0, M=1), race/ethnicity coded 0/1 as a series of dummy variables (Hispanic/Latino, White non-Hispanic/Latino, Black non-Hispanic/Latino, Multi-racial/Other non-Hispanic/Latino), IEP status (No IEP=0, IEP=1), English/Spanish *preLAS* language proficiency level assessed at fall of pre-k (4-level variable; *preLAS* levels 1-5, with levels 2 and 3 combined due to small sample sizes), and family income (Category Two=0, Category One=1)

2) Pre-k program/classroom characteristics – provider type (private setting=0, public school system=1) and class size
3) Classroom quality – Pre-k classroom quality (pre-K CLASS total) and subsequent K–2nd-grade classroom quality coded as a time-varying covariate (CLASS K-3 categorized as a three-level variable based on domain scores: high=ES & CO≥5 & IS≥2.5; medium=does not meet high criteria, ES & CO ≥4 & IS≥2; and low=ES<4, CO<4, or IS<2).

All covariates were crossed with Time and Time-squared to test whether child/family characteristics, pre-k program/classroom characteristics, or classroom quality predicted any differences in linear or quadratic growth effects for outcomes assessed from pre-k through second grade. For outcomes assessed beginning in first grade, only the differences in Time effects were tested given the limited number of outcome data points.

Two more sets of models were tested, following a similar algorithm as above, for the DLL subsample to examine results for parallel versions of outcome measures in English and in Spanish. (SSiS Social Skills and Problem Behaviors were excluded since these were based on teacher ratings rather than individual child assessments.) For analyses involving English outcome measures for the full sample or the DLL subsample, the English language proficiency level at pre-k entry was used; for analyses involving Spanish outcome measures for the DLL subsample, the Spanish language proficiency level at pre-k entry was used. In addition, all analyses for the DLL subsample excluded two predictors because of sample size issues; nearly all children in the subsample were Hispanic/Latino so race/ethnicity was excluded and only three children had an IEP, so IEP status was excluded.

Analyses involved grand mean centering, mean-centered covariates, accounted for missing data on covariates, and adjusted for multiple comparisons. The grand mean centering of the covariates meant that coefficients for time and time-squared described rates of change over time for the whole sample and took into account the main effects of the predictors. The HLMs analyzed all observed data, and assumed data were missing at random (i.e., they account for missing data on the outcome variables to the extent that the reason the data are missing is not causally related to the child's outcome scores[xvii]). Econometric methods were used to estimate missing values; mean values were assigned to replace missing data and missing data dummy variables were included in the analyses in order to describe associations between that predictor and outcome for all cases without missing data. Benjamini-Hochberg adjustments were applied to all the estimated coefficients involving time (time, interactions with time, time squared, and interactions with time squared) after all models were completed to account for multiple tests and to correct for the potential false discovery rate. Post-hoc analyses were conducted for the remaining significant interactions after the adjustments, and those estimates were used in the interpretation of results and for graphing the significant interactions. Tables do not include p-values for the intercept and covariates because they were not included in the Benjamini-Hochberg adjustments.

Full Sample 3-level HLM for outcomes collected from pre-k through grade 2: Level 1 – individual quadratic growth curve, Level 2 – child, Level 3 – pre-k classroom

$$
\begin{aligned}
DV = {} & B_0 + B_1\,timePT + B_2 timePT^2 + B_3\,boy + B_4\,PreLaspk{=}1 + B_5\,Prelaspk{=}2or3 \\
& + B_6\,Prelaspk{=}4 + B_7\,Black + B_8\,Hispanic + B_9\,Multi + B_{10}\,IEP + B_{11}\,poor \\
& + B_{12}\,CLASSPK + B_{13}\,public + B_{14}\,clsizePk + B_{15}\,CLASS\text{-}hi + B_{16}\,CLASS\text{-}mod \\
& + B_{17}\,boy{*}timept + B_{18}\,PreLaspk{=}1{*}timept + B_{19}\,PreLaspk{=}2/3{*}timept \\
& + B_{20}\,PreLaspk{=}4{*}timept + B_{21}\,Black{*}timept + B_{22}\,Hispanic{*}timept + B_{23}\,Mult{*}timept \\
& + B_{24}\,IEP{*}timept + B_{25}poor{*}timept + B_{26}\,CLASSPK{*}timept + B_{27}\,public{*}timept \\
& + B_{28}\,clsizePk{*}timept + B_{29}\,CLASS\text{-}hi{*}timept + B_{30}\,CLASS\text{-}mod{*}timept \\
& + B_{31}\,boy{*}timept^2 + B_{32}\,PreLaspk{=}1{*}timept^2 + B_{33}\,PreLaspk{=}2/3{*}timept^2 + \\
& + B_{34}\,PreLaspk{=}4{*}timept^2 + B_{35}Black{*}timept^2 + B_{36}Hispanic{*}timept^2 + B_{37}Mult{*}timept^2 \\
& + B_{38}\,IEP{*}timept^2 + B_{39}\,poor{*}timept^2 + B_{40}\,CLASSPK{*}timept^2 \\
& + B_{41}\,public{*}timept^2 + B_{42}\,CLASS\text{-}hi{*}timept + B_{43}\,CLASS\text{-}mod{*}timept \\
& + l_{0k} + l_{ik} + d_{0ik} + d_{1ik} + e_{ijk}
\end{aligned}
$$

For ith child in kth Pre-K class at jth timepoint, where l_{0k} is the random intercept for kth pre-k classroom, l_{0k} is the random slope and l_{1k} is the random intercept for kth pre-k classroom, d_{0ik} is the random slope and d_{1ik} is the random intercept for ith child in kth pre-k classroom, and e_{ijk} is the random residual for the jth assessment of ith child in kth pre-k classroom

Full Sample 3-level HLM for outcomes collected from grade 1 through grade 2: Level 1 – individual quadratic growth curve, Level 2 – child, Level 3 – pre-k classroom

$$
\begin{aligned}
DV = {} & B_0 + B_1\,timePT + B_2 timePT^2 + B_3\,boy + B_4\,PreLaspk{=}1 + B_5\,Prelaspk{=}2or3 \\
& + B_6\,Prelaspk{=}4 + B_7\,Black + B_8\,Hispanic + B_9\,Multi + B_{10}\,IEP + B_{11}\,poor \\
& + B_{12}\,CLASSPK + B_{13}\,public + B_{14}\,clsizePk + B_{15}\,CLASS\text{-}hi + B_{16}\,CLASS\text{-}mod \\
& + B_{17}\,boy{*}timept + B_{18}\,PreLaspk{=}1{*}timept + B_{19}\,PreLaspk{=}2/3{*}timept \\
& + B_{20}\,PreLaspk{=}4{*}timept + B_{21}\,Black{*}timept + B_{22}\,Hispanic{*}timept + B_{23}\,Mult{*}timept \\
& + B_{24}\,IEP{*}timept + B_{25}poor{*}timept + B_{26}\,CLASSPK{*}timept + B_{27}\,public{*}timept \\
& + B_{28}\,clsizePk{*}timept + B_{29}\,CLASS\text{-}hi{*}timept + B_{30}\,CLASS\text{-}mod{*}timept \\
& + l_{0k} + d_{0ik} + e_{ijk}
\end{aligned}
$$

For ith child in kth pre-k class at jth timepoint

DLL Subsample 3-level HLM for outcomes collected from pre-k through grade 2: Level 1 – individual quadratic growth curve, Level 2 – child, Level 3 – pre-k classroom

$$
\begin{aligned}
DV = {} & B_0 + B_1\,time + B_2 time^2 + B_3\,boy + B_4\,PreLaspk\text{-}1 + B_5\,Prelaspk\text{-}2or3 \\
& + B_6\,Prelaspk{=}4 + B_7\,Black + B_8\,poor + B_0\,public + B_{10}\,clsizePk \\
& + B_{11}\,PK\ CLASS\ TOTAL + B_{12}\,Annual\ CLASS{=}hi + B_{13}\,Annual\ CLASS\text{-}mod \\
& + B_{14}\,boy{*}time + B_{15}\,PreLaspk{=}1{*}time + B_{16}\,PreLaspk{=}2/3{*}time \\
& + B_{17}\,PreLaspk{=}4{*}time + B_{18}poor{*}time + B_{19}\,public{*}time + B_{20}\,clsizePk{*}time \\
& + B_{21}\,PKCLASStot{*}time + B_{22}\,CLASS\text{-}hi{*}time + B_{23}\,CLASS\text{-}mod{*}time \\
& + B_{24}\,boy{*}time^2 + B_{25}\,PreLaspk{=}1{*}time^2 + B_{26}\,PreLaspk{=}2/3{*}time^2 + \\
& + B_{27}\,PreLaspk{=}4{*}time^2 + B_{28}\,poor{*}time^2 \quad B_{29}\,public{*}time^2 + B_{30}\,clsizePk{*}time^2 \\
& + B_{31}\,PKCLASStot{*}time^2 + B_{32}\,CLASS\text{-}hi{*}time^2 + B_{33}\,CLASS\text{-}mod{*}time^2 \\
& + l_{0k} + l_{ik} + d_{0ik} + d_{1ik} + e_{ijk}
\end{aligned}
$$

For ith child in kth pre-k class at jth timepoint, where l_{0k} is the random intercept for kth pre-k classroom, l_{0k} is the random slope and l_{1k} is the random intercept for kth pre-k classroom, d_{0ik} is the random slope and d_{1ik} is the random intercept for ith child in kth pre-k classroom, and e_{ijk} is the random residual for the jth assessment of ith child in kth pre-k classroom

DLL Subsample 3-level HLM for outcomes collected from grade 1 through grade 2: Level 1 – individual quadratic growth curve, Level 2 – child, Level 3 – pre-k classroom

$$
\begin{aligned}
DV = {} & B_0 + B_1\,time + B_2 time^2 + B_3\,boy + B_4\,PreLaspk\text{-}1 + B_5\,Prelaspk\text{-}2or3 \\
& + B_6\,Prelaspk{=}4 + B_7\,Black + B_8\,poor + B_0\,public + B_{10}\,clsizePk \\
& + B_{11}\,PK\ CLASS\ TOTAL + B_{12}\,Annual\ CLASS{=}hi + B_{13}\,Annual\ CLASS\text{-}mod \\
& + B_{14}\,boy{*}time + B_{15}\,PreLaspk{=}1{*}time + B_{16}\,PreLaspk{=}2/3{*}time \\
& + B_{17}\,PreLaspk{=}4{*}time + B_{18}poor{*}time + B_{19}\,public{*}time + B_{20}\,clsizePk{*}time \\
& + B_{21}\,PKCLASStot{*}time + B_{22}\,CLASS\text{-}hi{*}time + B_{23}\,CLASS\text{-}mod{*}time \\
& + B_{24}\,boy{*}time^2 + B_{25}\,PreLaspk{=}1{*}time + B_{26}\,PreLaspk{=}2/3{*}time^2 + \\
& + l_{ok} + d_{0ik} + e_{ijk}
\end{aligned}
$$

For ith child in kth pre-k class at jth timepoint

References

[i] Bright from the Start: Georgia Department of Early Care and Learning. (2018). *Georgia Early Learning and Development Standards (GELDS)*. Retrieved from http://www.gelds.decal.ga.gov/Resources.aspx

[ii] Georgia Department of Education (2015). *Georgia's Performance Standards for Kindergarten*. Retrieved from https://www.georgiastandards.org/Georgia-Standards/Pages/default.aspx

[iii] Bright from the Start: Georgia Department of Early Care and Learning and Georgia State University. (2013). *Work Sampling Online training manual.* Retrieved from http://decal.ga.gov/documents/attachments/WSO_Training_Manual2014.pdf

[iv] Meisels, S. J., Marsden D. B., Dichtelmiller, M. K., Dorfman, A. B., & Jablon, J. R. (2001). *The Work Sampling System 4th Edition*. San Antonio, TX: Pearson

[v] Bright from the Start: Georgia Department of Early Care and Learning. (2017). *2017–2018 School year pre-k providers' operating guidelines.* Retrieved from http://www.decal.ga.gov/documents/attachments/Guidelines.pdf

[vi] Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III Tests of Achievement*. Itasca, IL: The Riverside Publishing Company.

[vii] Muñoz-Sandoval, A. F., Woodcock, R. W., McGrew, K. S., & Mather, N. (2005). *Batería III Pruebas de aprovechamiento*. Itasca, IL: Riverside Publishing.

[viii] 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.

[ix] Gresham, F., & Elliott, S. (2008). *Social Skills Improvement System*. Minneapolis, MN: Pearson.

[x] Duncan, S. E., & De Avila, E. A. (1998). *PreLAS 2000*. Monterey, CA: CTB/McGraw-Hill.

[xi] Hallgren, K.A. (2012). Computing inter-rater reliability for observational data: An overview and tutorial. *Tutorials in Quantitative Methods for Psychology, 8*(1), 23–34.

[xii] 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.

[xiii] 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.

[xiv] Pianta, R., Vitiello, V., Whittaker, J., & Ruzek, E. (2018, March). *Understanding the Effects of Classroom Processes on Child Outcomes in Pre-Kindergarten.* Paper presented at the Conference of the Society for Research on Educational Effectiveness, Washington, DC.

[xv] Duncan, G. J., & Magnuson, K. (2013). Investing in preschool programs. *Journal of Economic Perspectives*, *27*(2), 109-32.

[xvi] Phillips, D. A., Lipsey, M. W., Dodge, K. A., Haskins, R., Bassok, D., Burchinal, M. R., . . . Weiland, C. (2017). *The Current State of Scientific Knowledge on Pre-Kindergarten Effects.* Retrieved from https://www.brookings.edu/research/puzzling-it-out-the-current-state-of-scientific-knowledge-on-pre-kindergarten-effects/

[xvii] Widaman, K.F. (2006). Missing data: What to do with or without them. In K. McCartney, M. R. Burchinal & K. L. Bub (Eds.), *Monographs of the Society of Research in Child Development, 71(3)*, 42-64. https://doi.org/10.1111/j.1540-5834.2006.00404.x