NIEER
NATIONAL INSTITUTE FOR EARLY EDUCATION RESEARCH

Cultivate Learning
UNIVERSITY of WASHINGTON

YEAR 2 REPORT: SEATTLE PRE-K PROGRAM EVALUATION

October 2017

Milagros Nores,Ph.D., Steve Barnett,Ph.D., Gail Joseph,Ph.D., Sara Stull,Ph.D. Kwanghee Jung, Ph.D. &., Janet S. Soderberg, Ph.D., The National Institute for Early Education Research & Cultivate Learning

About the Authors

Milagros Nores, Ph.D. Dr. Nores is a Co-Director of Research at The National Institute for Early Education Research (NIEER) at Rutgers University. Dr. Nores conducts research at NIEER on issues related to early childhood policy, programs, and evaluation, both nationally and internationally. She is also on staff with the Center for Enhancing Early Learning Outcomes (CEELO), a federally funded comprehensive center that provides technical assistance to state agencies around early childhood.

W. Steve Barnett, Ph.D. Dr. Barnett is a Senior Co-Director of the National Institute for Early Education Research (NIEER) and a Board of Governors Professor at Rutgers University. He is also Principal Investigator of the Center for Enhancing Early Learning Outcomes (CEELO). His research includes studies of the economics of early care and education including costs and benefits, the long-term effects of preschool programs on children's learning and development, the economics of human development, practical policies for translating research findings into effective public investments and the distribution of educational opportunities.

Gail Joseph, Ph.D. Dr. Joseph is an associate professor in the area of Educational Psychology and Early Childhood and Family Studies at the University of Washington. She teaches courses, advises students, provides service and conducts research on topics related to early care and education. She is the Founding Director of the EarlyEdU Alliance and principal investigator and Director of Cultivate Learning at the University of Washington's (previously known as the Childcare Quality and Early Learning Center for Research and Professional Development, CQEL).

Sara Stull, Ph.D. Dr. Stull is a Research Associate at Cultivate Learning at the University of Washington's. She leads CL's work on the SPP Evaluation Study and conducts research at CL on topics related to early childhood education and assessment. She has contributed to CL's research on kindergarten entry assessment and QRIS evaluation.

Kwanghee Jung, Ph.D. Dr. Jung is an Assistant Professor at The National Institute for Early Education Research (NIEER) at Rutgers University. Her expertise is in quantitative data analysis and the effect of participation in child care and early education on children’s learning and development.

Janet S. Soderberg, Ph.D. Dr. Soderberg is the Director of Research and Evaluation at Cultivate Learning at the University of Washington. Dr. Soderberg's research includes exploration of the association between classroom quality and children’s development, QRIS program evaluation, refinement and support of kindergarten entrance assessments, and dissemination of research. Her interests include child development, assessment, childcare quality, and multi-systems alignment.

Grateful acknowledgment is made to Erica Johnson and all the support from Seattle's Preschool Program and Joelle Gruber from Third Sector Intelligence for their support on this project. The authors are also grateful to Ruiya Zhang, Ran Guo and Zijia Li with their assistance in producing this report.

Correspondence regarding this report should be addressed to Milagros Nores at the National Institute for Early Education Research. Email: mnores@nieer.org.

Suggested citation: Nores, M., Barnett, W.S., Joseph, G., Stull, & Soderberg, J.S. (2017). Year 2 report: Seattle Pre-k program evaluation. New Brunswick, NJ: National Institute for Early Education Research & Seattle, WA: Cultivate Learning.

Table of Contents

Summary of Findings and Recommendations

Throughout the academic year of 2016-2017, the National Institute for Early Education Research at Rutgers University and Cultivate Learning, at the University of Washington conducted an evaluation of the second year of the demonstration phase of the Seattle Preschool program. This evaluation: (1) assessed preschool quality using two observational measures, (2) measured children's gains in receptive vocabulary, literacy, math and executive functions, and (3) compared gains of SPP attendees to those of a non-equivalent comparison group to estimate SPP impacts on children's learning and development. The non-equivalent comparison, while not ideal, was the best available approach to provide some indication of SPP's impact on children during this initial development period.

Overall, SPP in its demonstration phase has improved quality while expanding in size. In 2016-17 the program had almost double the number of classrooms yet both new and "old" classrooms exceeded last year's average level of quality. SPP provides caring and nurturing environments for children. Quality compares favorability to that of other well-known city programs, and instructional quality now exceeds an accepted threshold for effectiveness. Consistent with this, we also find evidence of improvements in children's learning.

Continuing this trajectory of quality improvement with growth in access will take SPP along the right path to achieve the goals of the demonstration phase. There is room for continued improvement along this path. Professional development and coaching could usefully focus on strengthening personal care routines and learning activities as measured by the ECERS-3 and instructional support as measured by the CLASS. To be more specific, important areas for improvement likely include increasing the amount of rich content, increasing integration across content areas, reducing transition time, and supporting metacognition in settings that provide high levels of individualization and choice. This report provides rich content that should support such work.

This general summary above is based on a much more detailed and specific analysis and report that answers 6 specific questions. We briefly summarize the answers below as a guide to what can be found in the full report.

1. Who enrolled in SPP classrooms in 2016–17, and how do they compare demographically to children in Seattle generally?

SPP children were similar in their personal and family background characteristics to those in Seattle Public Schools. They did not differ from the general public-school population with respect to gender, language, or percentage below the federal poverty level. SPP children in the study were slightly more likely to be Black or Multi-Racial than public school children generally.

2.What was the quality of children's SPP classroom experiences in 2016–17 and did it improve over the prior year.

Even with the growth from 14 classrooms in 2015-16 to 32 classrooms in 2016-17, the program improved on the Early Childhood Environmental Rating Scale – Third edition (ECERS -3) and the Classroom Assessment Scoring System (CLASS). All SPP classrooms had higher quality than classrooms attended by the comparison group.

The average ECERS 3 rating improved from 3.57 to 3.89 (on a 7-point scale). CLASS scores maintained already high levels on Emotional Support (going from 6.14 to 6.29) and Classroom Organization (from 5.67 to 5.55). CLASS Instructional Support score improved from 2.65 to 3.06 (also on a 7-point scale).

3. How does quality vary within SPP, and do children from different backgrounds experience different quality?

Quality was looked at across agencies, class size, and various child demographics. We found minor variation in quality, but differences by child background were not statistically significant with a few exceptions.

With program expansion from 14 to 32 classrooms any increase in quality could be solely due to incoming classrooms being of better quality. We find that this is not the case in ECERS-3 and that for the CLASS there are minimal advantages of new classrooms.

4. How did the learning of children enrolled in SPP classrooms progress in 2016–17, and how did it vary with classroom quality?

Minority children made the largest gains compared to White non-Hispanic children. Dual language children had larger gains in vocabulary, literacy and math. No differences were found by income, though there is a trend toward greater gains for children in poverty.

In terms of classroom characteristics, the differences observed in classroom size did not relate to children's performance, while there were a few associations detected between ECERS-3 and CLASS classroom organization scores and children's literacy outcomes. However, the small number of classrooms provides little statistical power to detect relationships between classroom characteristics and child outcomes.

To measure program impact we compared children's performance in SPP with that of children in a comparison group. We found evidence suggesting SPP had positive effects on vocabulary, literacy, math and one measure of executive function (DCCS), although these were mostly not statistically significant. A statistically significant negative SPP effect was found for the other measure of executive function (Peg Tapping).

5. To what extent are children's learning gains moderated by other learning activities, particularly parent activities and prior center-based care?

SPP parents reported higher levels of connection to preschool and communication with the teachers. Both parent outcomes are positively associated with some outcomes and moderated the impact of the program.

6. What activities do children engage in?

We found an improvement in the quality of staff-child interactions, and that a higher percentage of classrooms had strong interactions this year compared to last year. This improvement occurred even with the program increasing in size from 14 classrooms in 2015–16 to 32 classrooms in 2016–17. SPP classrooms improved on 70 percent of the 47 indicators focused on interactions with children.

Introduction

The City of Seattle is currently in the second year of a four-year demonstration phase for its Seattle Preschool Program. The program was established by voter approval on November 4, 2014 of a four-year, $58 million property tax levy to provide "accessible high-quality preschool services for Seattle children designed to improve their readiness for school and to support their subsequent academic achievement." The city of Seattle's Department of Education and Early Learning (DEEL) launched SPP in the 2015–16 school year and expanded it in the 2016–17 school year.

The four-year demonstration phase of SPP has three purposes. The first is to demonstrate that the approved structure is viable. The second is to develop a community infrastructure to improve the quality of preschool programs. The third is to create norms and a process to support continuous quality improvement (CQI) through evaluation. Results from evaluation during the demonstration phase will inform improvements in these efforts.

Before commencement of this demonstration phase, the evaluation team conducted a thorough review of the research on evaluation, supplemented with interviews of key leaders in program design and improvement. Based on this review the team recommended, among other efforts, an impact evaluation that collects information on students' learning and development.

This report presents second year (2016–17) findings from the impact evaluation, focusing on classroom quality and children's learning. We report for SPP basic statistics that describe: the children served, children's learning and development including average gains during the year, and program quality. As context, we report similar information from other preschool studies including the Head Start Family and Child Experiences Survey (FACES). We report findings for subgroups of students and classrooms as well as the full sample. To estimate program impact we compare learning gains of SPP children to those of other children who did not attend SPP. Finally, to inform those developing and implementing the program we investigate relationships between SPP children's learning gains and their classroom experiences including observed quality.

As with all evaluations, this one has limitations that must be acknowledged. The second year of the impact evaluation employs a non-equivalent comparison group design to estimate program impacts on children's learning and development. Comparison groups were constructed using data on children from DEEL waiting lists for SPP and other children enrolled in centers that waiting list children attended in 2016–17. This design relies on statistical controls to adjust for differences between the groups other than SPP participation. Although not ideal, this was the best available approach to provide some indication of SPP's impact on children during this initial development period. As the program is far from fully established, greater weight is given during the demonstration period to establishing that the program and its infrastructure are being developed to meet expectations for program performance as the system matures.

Study Methods

The SPP evaluation study is a multi-year, multi-site study that combines various designs to provide a comprehensive assessment of the program's quality and its impact on children over time. The second year of the study included collection of child, family and classroom information to address the following six questions:

1. Who were the children enrolled in SPP classrooms in 2016–17, and how do they compare demographically to children in Seattle more generally?
2. What was the observed quality of children's SPP classroom experiences in 2016–17? Did SPP quality improve over the prior year?
3. How does quality vary within SPP and do children from different backgrounds experience different quality?
4. How did the learning of children enrolled in SPP classrooms progress in 2016–17, and how did it vary with classroom quality?
5. To what extent are children's learning gains moderated by other learning activities, particularly parent activities and prior center-based care and education?
6. What activities do children engage in, and is there scope for their interests and active participation?

The SPP evaluation was framed to understand the effects of SPP on children's learning and development. In Year 1, the research team measured learning and development at the beginning and at the end of the year. In Year 2, the research team repeated this process, and also recruited a non-equivalent comparison group that is composed of children in the waiting list for SPP together with children attending centers were some waiting list children ended up enrolled. Measures and procedures are described below. Children were assessed in the Fall of 2016, and assessed again at the end of the school year. Moreover, classroom observations of classroom practices were conducted to assess overall quality, teacher-child interactions, and engagement. Classroom observations were conducted between the months of February through March. Quality was assessed using observation protocols widely established in the field. Figure 1 (below) reports the data collection timeline for the school year of 2016–17.

Figure 1. Data Collection Timeline

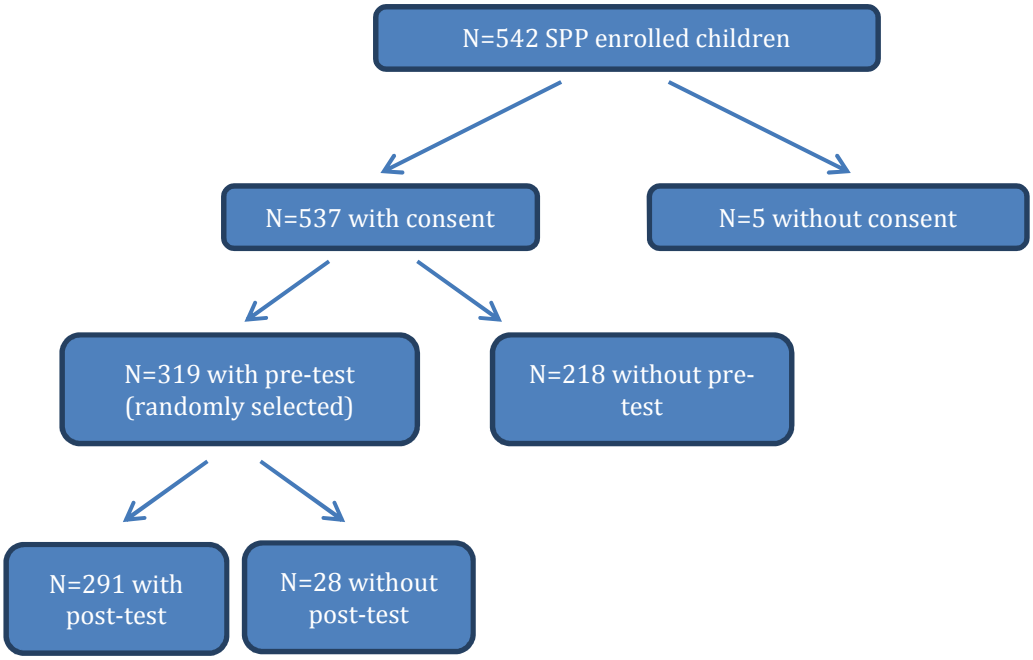
2016	
September	• Training for data collectors • Initial SPP site information gathered
October	• Parent consent form distribution • Fall assessment visit scheduling • Fall child assessment visits begin
November	• Fall child assessment visits continue (only one site required that we go into December for child assessments)
2017	
January	• Communications to directors to discuss classroom observations (CLASS & ECERS-3)
February	• Unannounced CLASS & ECERS-3 observations (February and March)
March	• Roll out of parent survey
April - June	• Spring assessment visit scheduling (early April) • Spring child assessment visits

Sample

In the AY 2016–17 the research team assessed 291 children in 32 SPP classrooms at pre- and post-test, and 137 children not enrolled in SPP classrooms. To recruit children in SPP, we

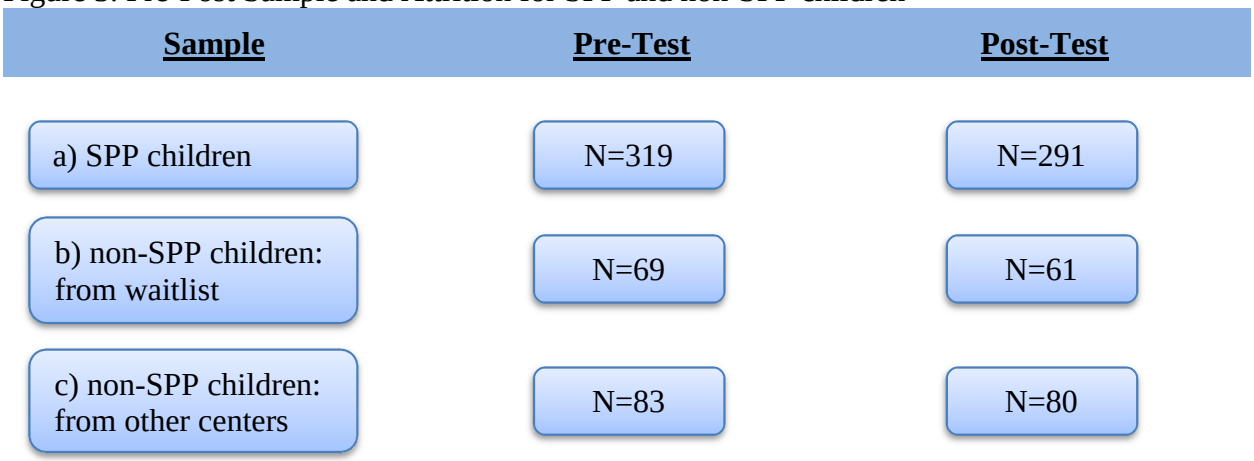
distributed consent forms across all classrooms. Of the parents of the 542 children enrolled in these classrooms, 537 consented to participate in the study. Some of these children were grandfathered into the program, meaning they were not admitted through the DEEL list but rather because they had been enrolled in the program a previous year or had siblings enrolled (agency selected children). We focused on recruiting children that had been enrolled through the SPP list, unless all children in a classroom had been selected by the Agency. We selected (randomly in classrooms where it was feasible) 319 children from SPP classrooms. Out of these, seven children declined participation, and 291 were assessed at post-test on at least one measure. Figure 2 below shows the study attrition tree. Seven children required language accommodations.

Figure 2. Pre-Post Sample Attrition Tree



We also recruited 152 non-SPP children, both, from the waiting list and from centers were waiting list children were enrolled. Of the group of non-SPP children, 69 were recruited from the waiting list and 61 of these followed at post-test. In addition, another 83 children were recruited from centers where waiting list children attended, and 80 of these followed at post-test. Figure 3 below summarizes all three groups of children recruited for the study.

Figure 3. Pre-Post Sample and Attrition for SPP and non-SPP children



In addition, we conducted classroom observations on the 32 SPP and 7 non-SPP classrooms from which we drew children. SPP Classroom characteristics are described in Table 2. Most classrooms in SPP in Year 2 used either Creative Curriculum or HighScope Curriculum, they reported an average class size of 18, and they were distributed across ten agencies, with about three classrooms per agency.

Table 1. SPP Classroom characteristics, N=32

Classroom characteristic		Frequency or Mean (SD[1])
Curriculum	Creative	12
	HighScope	20
Class Size		16.53 (2.71)*
Agencies		10
Average No. Classrooms per Agency		3.20 (2.20)

* Number of children in classroom as reported by director/roster in October/November.

Measures

Measures on Children

The *Peabody Picture Vocabulary Test—Fourth Edition (PPVT-IV;* Dunn & Dunn, 2007) is a 228-item test of receptive vocabulary in standard English. The PPVT is predictive of general cognitive abilities and is a direct measure of vocabulary size. The rank order of item difficulties is highly correlated with the frequency with which words are used in spoken and written language. The test is adaptive (to avoid floor and ceiling problems), establishing a floor below which the child is assumed to know all the answers and a ceiling above which the child is assumed to know none of the answers and can be used with population ages 2.5 and above. The

[1] SD stands for standard deviation, which is a measure of variation in the data. That is, it measures how close together or spread apart the classrooms are relative to the mean. The larger the value, the farther apart from the mean classrooms are, and the smaller the value, the closer to the mean classrooms are, in a specific indicator, such as classroom size.

test is reliable based on reported split-half reliabilities or test-retest reliabilities. The PPVT has shown concurrent validity (e.g., Qi, Kaiser, Milan, & Hancock, 2006) and the results of these tests are found to be strongly correlated with school success (Blair & Razza, 2007; Early, et al., 2007).

The *Woodcock-Johnson Psycho-Educational Battery—Third Edition (WJ-III;* Woodcock, McGrew, Mather, & Schrank, 2001) includes multiple subtests. Only the *Applied Problems* and *Letter-Word Identification* subtests were used in this study. WJ-III was normed on a stratified random sample of 6,359 English-speaking subjects in the United States. Like the PPVT, the WJ is also an adaptive test used with populations above the age of 3. Correlations of the WJ with other tests of cognitive ability and achievement are reported to range from 0.60 to 0.70. This measure has been used in numerous large-scale preschool studies (e.g., Early, et al., 2007; Wong, Cook, Barnett, & Jung, 2008).

Dimensional Change Card Sort Task (DCCS; Zelazo, 2006). This task engages reverse categorization where children must sort a set of cards based on different sorting criteria given by the examiner. Generally, the test assesses attention-shifting. Scores on the DCCS reflect a pass/fail system on each of three levels of increasing difficulty. Raw scores range between 0 and 3, where a score of 0 means a child did not pass the first level which includes a color sorting task. At this first level, children are tasked with sorting two objects by color into a corresponding labeled box. A score of 1 means a child passed the color sort but failed the shape sort, which is the subsequent task and asks children to ignore color and instead sort objects by their shape. A score of 2 means a child passed shape sort but failed advance trials. Lastly, a score of 3 means the child passed advance trials, which ask children to ignore color or shape by adding a border to cards to indicate which attribute to sort by. There are no standard score equivalents. However, in a study of test-retest reliability, means by age for children age 48 months or younger were 1.14 for 48–50 months they were 1.33, for 51–53 months they were 1.42, and for 54–56 months they were 1.58 (Meador et al., 2013).

Peg Tapping Test (PT; Diamond & Taylor, 1996). In this game, children are asked to tap a peg twice when the experimenter taps once and vice versa. The task requires children to inhibit a natural tendency to mimic the experimenter while remembering the rule for the correct response. Sixteen trials are conducted with 8 one-tap and 8 two-tap trials in random sequence. The task requires both the ability to hold two things in mind—the rule to tap once when experimenter taps twice and the rule to tap twice when experimenter taps once, and the ability to exercise inhibitory control over one's proponent behavior, the natural tendency to mimic what the experimenter does. Common errors include: (1) complying with only one of the two rules, (2) tapping many times regardless of what the experimenter did, and (3) doing the same thing as the experimenter, rather than the opposite. The final score for Peg Tapping is a sum of all the 16 items that comprise the test. Again, while there are no standard score equivalents, in a study of test-retest reliability, means by age for children age 48 months or younger were 4.05, for 48–50 months they were 4.57, for 51–53 months they were 6.02, and for 54–56 months they were 7.87 (Meador et al., 2013).

We also conducted family surveys asking families to provide information regarding the following:

- Demographics of the child and family such as family income, education, employment status, race/ethnicity, and languages spoken at home,
- Learning activities in the home, and other types of care and education the child may receive outside the home

- Family perceptions of early education or child care programs, and family perspectives or the benefits of SPP including impacts on their child's learning and development

Parental response rate was high (81.9%). Consequently, we have folded information collected on families into respective analyses, to complement the information we received from DEEL on children and families, as well as provided additional tables in Appendix F.

Measures on Classrooms

Early Childhood Environment Rating Scale—Third Ed. (ECERS-3; Harms, Clifford & Cryer, 2014)

The ECERS-3 is an observation and rating instrument for preschool classrooms serving children aged three to five. The total ECERS-3 score represents an average of the scores on the 35 items under 6 domains. A rating scale between 1 and 7 is used, where a rating of 1 indicates inadequate quality, a rating of 3 indicates minimal quality, a rating of 5 indicates good quality, and a rating of 7 indicates excellent quality. The most updated notes for clarification[2] were utilized when scoring all classrooms in this sample. A general description of each of the 35 items on the ECERS-3 is provided in Table 2.

Table 2. ECERS-3 Subscale and Item Descriptions.

Subscale	Items	Description
Space for Furnishings	1. Indoor Space	Considers enough indoor space for children, staff, and basic furnishings for routines, play, and learning.
	2. Furnishings for care, play, and learning	Focuses on ample furniture for routine care, play, and learning, including convenient cubbies for individual use.
	3. Room arrangement for play and learning	Space is arranged so that classroom pathways generally do not interrupt play and supervision.
	4. Space for privacy	Considers an indoor space for privacy available and set up physically in the classroom to discourage interruptions.
	5. Child-related display	Focuses on appropriate materials displayed for children throughout the classroom, including simple pictures, posters, and artwork.
	6. Space for gross motor play	Gross motor area is spacious, generally safe, and easily accessible to children.
	7. Gross motor equipment	Equipment is age appropriate, accessible, and ample enough to interest every child.
Personal Care Routines	Meals/Snacks	Schedule and sanitary procedures are appropriate during meal times. Staff sit with children to encourage learning.
	Toileting/diapering	Proper sanitary procedures usually followed with pleasant supervision.
	Health practices	Proper sanitary procedures used consistently as needed, with a few lapses.
	Safety practices	Considers no more than 2 major safety hazards present indoors or outdoors.
Language and Literacy	Helping children expand vocabulary	Measures how frequent staff uses specific words for objects and actions and descriptive words as children experience routines and play.
	Encouraging children to use language	Assesses how frequent staff asks questions that children are interested in answering and that require longer answers. Includes many conversations during gross motor free play and routines.

[2] Published online at http://ersi.info/ecers3_notes.html in November, 2016.

	Staff use of books with children	Staff read appropriate books to children that relate to current classroom activities or themes, showing interest and enjoyment while doing so.
	Encouraging children's use of books	Many books are accessible and organized in a defined interest center.
	Becoming familiar with print	Focuses on how most visible print is combined with pictures, relates to current classroom topics, and shows a variety of words.
Learning Activities	Fine motor	Focuses on the accessibility for children of fine motor materials, including interlocking building materials, manipulatives, puzzles, and art materials.
	Art	Art materials, including drawing materials, paints, 3D objects, collage materials, and tools, must be accessible for children.
	Music and movement	Measures how many music materials and activities are accessible for children during free play.
	Blocks	Enough space, unit blocks and accessories from 3 different categories for 2-3 children to build at once.
	Dramatic play	Many and varied dramatic play materials, including dolls, furniture, play food and dress-up clothes must be accessible for children during free play.
	Nature/science	At least 15 nature/science materials, including living things, natural objects, factual books, tools, or sand/water must be accessible for children.
	Math materials and activities	At least 10 different appropriate math materials accessible, including materials to count/compare quantities, measure/compare sizes, and familiarize children with shapes.
	Math in daily events	Assess how staff encourages math learning as part of daily routines.
	Understanding written numbers	At least 3-5 different materials should be present in the classroom that shows children the meaning of print numbers.
	Promoting acceptance of diversity	At least 10 examples of diversity accessible, including books, displayed pictures and materials.
	Appropriate use of technology	All observed materials used are appropriate and limited to 10-15 minutes per child during the observation.
Interaction	Supervision of gross motor	Focuses on careful supervision in order to ensure children's safety.
	Individualized teaching and learning	Many activities observed are open- ended and most allow children to be successful.
	Staff-child interaction	Evaluates frequent positive staff- child interactions, with no long periods of no interaction.
	Peer interaction	Captures positive peer interactions during at least half of the observation.
	Discipline	Children appear to be aware of classroom rules, and generally follow them with reasonable amount of teacher control.
Program Structure	Transitions and waiting times	Classroom transitions are usually smooth and productively engaging.
	Free play	Free play takes place for 1 hour during observation, including some time indoors and some time outdoors (weather permitting).
	Whole - group activities for play and learning	Staff are responsive and flexible in ways that maximize child engagement during whole group activities.

Classroom Assessment Scoring System (CLASS; Pianta, La Paro, & Hamre, 2008)

The CLASS is an observational system that assesses classroom practices in preschool and kindergarten by measuring the interactions between students and adults. Observations consist of four-to-five 20-minute cycles, with 10-minute coding periods between each cycle.

Scores are assigned during various classroom activities, and then averaged across all cycles for an overall quality score. Interactions are measured through 10 dimensions, which are divided into three domains. The Emotional Support domain is measured by four dimensions: Positive Climate, Negative Climate, Teacher Sensitivity, and Regard for Student Perspectives. The Classroom Organization domain is measured by 3 dimensions: Productivity, Behavior Management, and Instructional Learning Formats. The Instructional Support domain is measured by 3 dimensions: Concept Development, Quality of Feedback, and Language Modeling. Each scale uses a 7-point Likert-type scale, for which a score of 1 or 2 indicates low quality, and a score of 6 or 7 indicates high quality. The CLASS domains and dimensions are outlined in Table 3 below.

Table 3. CLASS Domains and Dimension Descriptions.

Domain	Dimension	Description
Emotional Support	Positive Climate	Reflects the emotional connection between teachers and children and among children, and the warmth, respect, and enjoyment communicated by verbal and nonverbal interactions.
	Negative Climate	Reflects the overall level of expressed negativity in the classroom. The frequency, quality, and intensity of teacher and peer negativity are key to this dimension
	Teacher Sensitivity	Encompasses the teacher's awareness of and responsiveness to students' academic and emotional needs.
	Regard for Student Perspectives	Captures the degree to which the classroom activities and teacher's interactions with students place an emphasis on students' interests, motivations, and points of view and encourage student responsibility and autonomy.
Classroom Organization	Behavior Management	Encompasses the teacher's ability to provide clear behavior expectations and use effective methods to prevent and redirect misbehavior.
	Productivity	Considers how well the teacher manages instructional time and routines and provides activities for students so that they have the opportunity to be involved in learning activities.
	Instructional Learning Formats	Focuses on the ways in which teachers maximize students' interest, engagement, and abilities to learn from lessons and activities.
Instructional Support	Concept Development	Measures the teacher's use of instructional discussions and activities to promote students' higher-order thinking skills and cognition and the teacher's focus on understanding rather than on rote instruction.
	Quality of Feedback	Assesses the degree to which the teacher provides feedback that expands learning and understanding and encourages continued participation.
	Language Modeling	Captures the effectiveness and amount of teacher's use of language-stimulation and language-facilitation techniques.

Procedures

Data collection processes were led by Cultivate Learning at the University of Washington. The center hired and trained data collectors on the child standardized assessment and classroom

observation measures. Data collectors received a two-day training on the measures for child assessments. Following the training, data collectors were successfully shadowed by expert staff on two iterations of the assessments for reliability. After two iterations of assessments, each data collector achieved 100% reliability.

Trained and reliable observers are necessary for observations of classroom quality. Initial training was provided in administering the observation protocol that includes the ECERS-3 and the CLASS for preschool classrooms. Training took place in separate full-day workshops. ECERS-3 observers were trained by an ECERS-3 certified trainer and met the ERSI[3] reliability requirements for observer certification. The trainee must complete three observations with the trainer with an average of 85% or above exact matches or one-away from the true score. All data collectors met the ECERS-3 reliability requirements with agreement percentages ranging between 80–89%. CLASS observers were trained by a CLASS certified trainer and met the Teachstone reliability requirements for observer certification. All data collectors met CLASS reliability[4] requirements with agreement percentages ranging between 93–100%. All assessment and observation score sheets were cleaned and entered at UW by trained staff. Language accommodations were made as necessary in the requested language (N=5). Assessment procedures integrated culturally sensitive attitudes, knowledge, interview skills, intervention strategies and evaluation practices specifically informed by the age of the child.

Results

We have organized this section by research question, addressing each individually through a combination of descriptive and statistical analyses. Analyses draw from the sample classrooms, the sample of children in SPP classrooms, and the sample of non-SPP children described earlier, respectively.

1. Who were the children enrolled in SPP classrooms in 2016–17, and how do they compare demographically to children in Seattle more generally

Children's demographics[5] are summarized in Table 4, below, in comparison to children enrolled in Seattle Public Schools (rather than children in Seattle more generally, as these children more likely encompass the SPP program target population). Children in the sample were mostly 4-year-olds (83.8%) and predominantly from English-speaking households (64.7%), with 15.8% speaking other languages, including Vietnamese, Amharic, Mandarin, Somali, and Oromo (among others). About 67.4% of the children were under 300% of the Federal Poverty Level (FPL). Children had significant variation across parent-reported race and ethnicity, with the four major groups being White (20.6%), Black (23.8%), Asian (16.6%), Hispanic (7.5%), or Multiracial/ethnic (28.1%).

[3] ERSI is the company that sells ECERS-3 products. More information about the tool, as well as reliability guidelines, can be found at http://www.ersi.info/

[4] Teachstone is the company that sells CLASS products and manages/sells CLASS observer trainings, certifications etc. All training activity is monitored and reported to them. http://www.teachstone.com/about-teachstone/.

[5] Demographics are based primarily on parental survey responses, and for missing cases, on DEEL provided demographics. Race/ethnicity demographics between parental reports and DEEL differed for 19% of the sample.

Table 4. Child demographics for SPP children relative to children in Seattle Public Schools

Child Characteristics	SPP Children 2016–17 N	SPP Children 2016–17 %	Seattle Public Schools
Gender			
Female	141	48.5%	51.5%[a]
Male	150	51.6%	48.5%[a]
Age at Pre-Test			
3-Year-Olds	47	16.2%	-
4-Year-Olds	244	83.8%	-
Primary Language			
English	199	68.4%	65.0%[b]
Spanish	3	1.0%	7.0%[b]
Vietnamese	12	4.1%	3.0%[b]
Amharic	7	2.4%	<1.0%[b]
Chinese-Mandarin/Cantonese	9	3.1%	3.0%[b]
Somali	4	1.4%	4.0%[b]
Oromo	2	0.7%	<1.0%[b]
Other	9	3.1%	-
Unknown	46	15.8%	-
Income			
20,000 or Less	43	14.8%	
21,000-40,000	54	18.6%	
41,000-60,000	42	14.4%	
61,000-80,000	44	15.1%	
81,000 or more	59	20.3%	
Unknown	49	16.8%	
FPL Percentage			
Less than 100%	61	21.0%	38.9%[a,c]
100 – 199%	68	23.4%	
200 – 299%	67	23.0%	-
≥ 300%	95	32.7%	-
IEP/IFSP			
Yes	16	5.5%	-
Unkown	38	13.1%	-
Race/Ethnicity			
White	62	21.3%	45.6%[a]
Black	67	23.0%	16.4%[a]
Asian	48	16.5%	15.8%[a]
Hispanic	23	7.9%	12.4%[a]
Multi-Racial	83	28.5%	8.5%[a]
Other	8	2.8%	1.3%[a]

[a]Seattle Public Schools as reported in http://www.seattleschools.org/district/district_quick_facts.
[b]Students attending Seattle Public Schools, as reported in Rivers (2016).
[c]Based on Free and Reduce Lunch which is for families <185% FPL.

Families in the SPP program that responded to the survey (82% of SPP families, see details in Appendix F) were also found to be mostly living with two parents (75%, whether both biological or not), with on average 3.3 years in residence in the current home, and with mothers having given birth to the child in the program at an average age of 31. Fifty percent of the

families own more than 50 books in their home, and only 11% reported using after care. The family survey also included questions on food fragility and the extent to which families found themselves sometimes or often experiencing concerns on food: 27% of families in SPP that answered the survey reported worrying food would run out, 18% reported that food would not last, 20% reported not being able to afford balanced meals, and 22% reported relying on only a few types of low-cost meals. Families in SPP reported being somewhat dependent on food stamps (24%), WIC (21%), Medicaid (28%) and Medicare (16%).

Table 5. Family dependency on welfare programs

	Food Stamps		WIC		Medicaid		Medicare	
	n	%	n	%	n	%	n	%
Yes	57	19.6	49	16.8	66	22.7	35	12.0
No	181	62.2	191	65.6	168	57.7	195	67.0
N/A	53	18.2	51	17.5	57	19.6	61	18.2
Total	291	100.0	291	100.0	291	100.0	291	100.0

Note. N/A = Don't know or Missing

2. What was the observed quality of children's SPP classroom experiences in 2016–17? Did SPP quality improve over the prior year?

Average ECERS-3 Results

A summary of ECERS-3 scores for all SPP classrooms are reported in Table 6 below. The table shows the mean scores, standard deviation, and the minimum and maximum scores, for all six ECERS-3 subscales and for overall scores. Classrooms scores for this year are compared with those of the previous year. Average overall ECERS-3 scores and subscale scores in 2017 were slightly higher (increase is equivalent to 0.70SD) despite the increase in the number of classrooms in the program. However, variation also increased, with lower and higher scoring classrooms present in the program.

Table 6. ECERS-3 Item, Subscale, and Overall Means and Ranges, 2016 & 2017

ECERS-3 Item and Subscales	Spring 2016 (N=14)				Spring 2107 (N=32)			
	Mean	(SD)	Min.	Max.	Mean	(SD)	Min.	Max.
Overall	**3.57**	**(0.46)**	**2.94**	**4.50**	**3.89**	**(0.55)**	**2.74**	**5.44**
Space and Furnishings	3.88	(0.55)	2.86	4.57	3.94	(0.61)	2.71	5.29
Personal Care Routines	3.14	(0.65)	1.75	4.25	3.41	(0.86)	1.50	5.50
Language and Literacy	3.47	(0.83)	2.40	5.20	3.93	(0.82)	2.40	6.00
Learning Activities	2.87	(0.56)	2.10	4.00	3.26	(0.57)	2.40	4.70
Interaction	4.49	(0.90)	3.20	5.80	4.99	(1.07)	2.40	6.80
Program Structure	4.43	(0.97)	2.67	6.00	4.67	(0.88)	3.00	6.33

Average CLASS Scores

Classrooms were also observed using the CLASS. The scores presented in Table 7 reflect overall means for all pre-K classrooms in the SPP program in the spring of 2016 and 2017. The table presents the mean scores, standard deviation, and the minimum and maximum scores for all three

CLASS domains. Two domains, emotional support and instructional support, evidence an increase in mean scores in 2017 relative to 2016 (increase is equivalent to 0.28SD and 0.58SD, respectively), and higher minimums and maximums. Classroom organization evidenced slightly lower scores (-0.16SD).

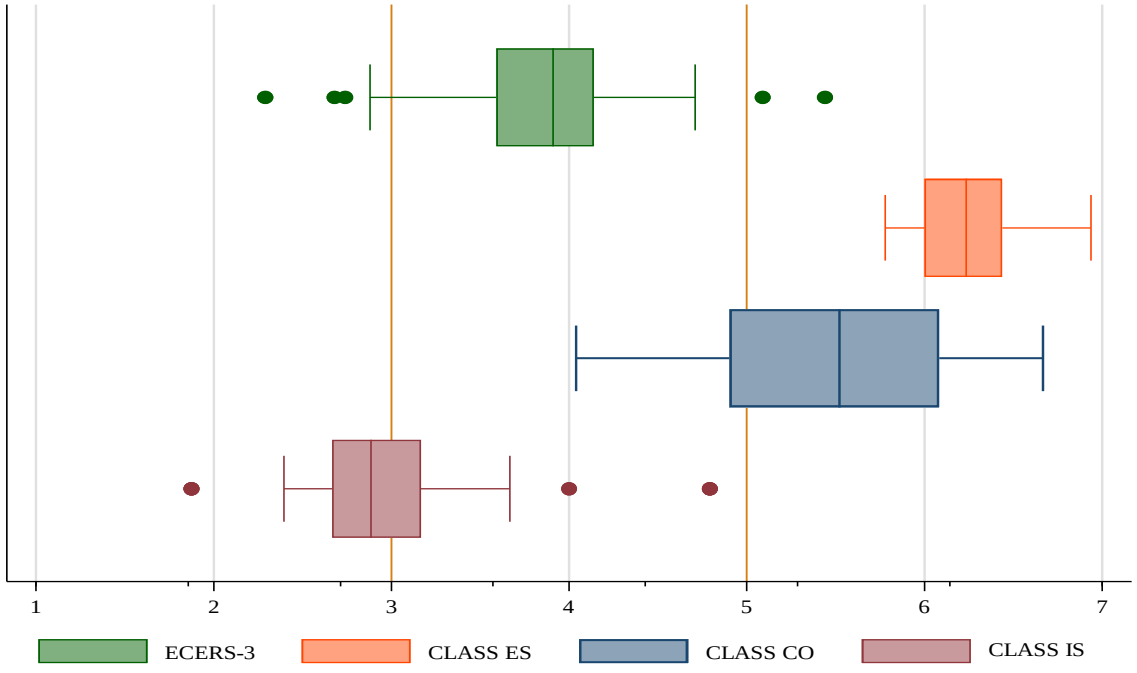
Table 7. CLASS Dimension and Domain Means and Ranges, 2016 & 2017

CLASS Dimensions and Domains	Spring 2016 (N=14)				Spring 2107 (N=32)			
	Mean	(SD)	Min.	Max.	Mean	(SD)	Min.	Max.
Emotional Support	6.14	(0.53)	4.88	6.81	6.29	(0.47)	5.19	7.00
Classroom Organization	5.67	(0.74)	4.17	6.58	5.55	(0.76)	3.42	6.83
Instructional Support	2.65	(0.71)	1.50	4.25	3.06	(0.88)	1.67	5.75

Distribution of Classroom Quality across Classrooms

The distribution of classroom quality as measure by ECERS-3[6] and CLASS domains are depicted in Figure 4.1, below. In the spring of 2017, classroom scored well below the good-quality threshold of 5 in the ECERS-3. On the other hand, classrooms score quite high on Emotional Supports, with most classrooms heavily concentrated around the mean score of 6.14. Classroom organization also had elevated scores, with a large portion of classrooms scoring above 5. Classrooms score lower on instructional support, with 75% of the classrooms scoring under 3.5. These patterns are consistent with patterns observed in the overall field.

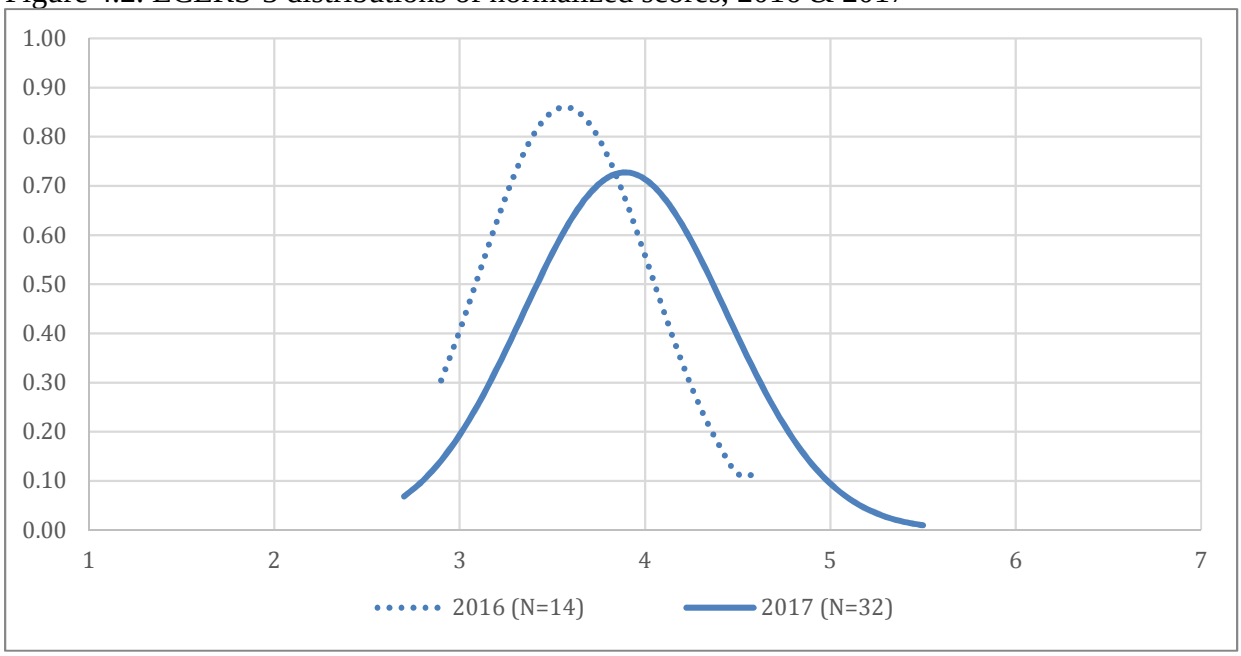
Figure 4.1. ECERS-3 and CLASS Domain distributions of scores as box plot, 2017



[6] ECERS-3 is a newer version of the widely used ECERS-R measure. Like the ECERS-R, quality in the ECERS-3 is considered minimal when the average or subscale scores is between 3 and 5, as is the case of SPP classrooms. Higher quality classrooms are expected to average a score between 5 and 7.

In addition, Figures 4.2 and 4.3 illustrate normalized distributions for ECERS-3 and CLASS dimensions for the spring of 2016 (dotted line) and 2017 (solid line). The ECERS-3 distribution of classrooms evidences a larger portion of classrooms scoring higher in the scale, relative to 2016.

Figure 4.2. ECERS-3 distributions of normalized scores, 2016 & 2017



There are also marked differences in the distributions of CLASS (Figure 4.3), in particular for the instructional support and emotional support domains. The 2017 classrooms in the SPP program reached higher IS scores and were spread out more, with a higher percentage of classrooms scoring over 3. Similarly, classrooms in the 2017 sample were concentrated at higher ES scores.

Figure 4.2. CLASS Domain distributions of normalized scores, 2016 & 2017

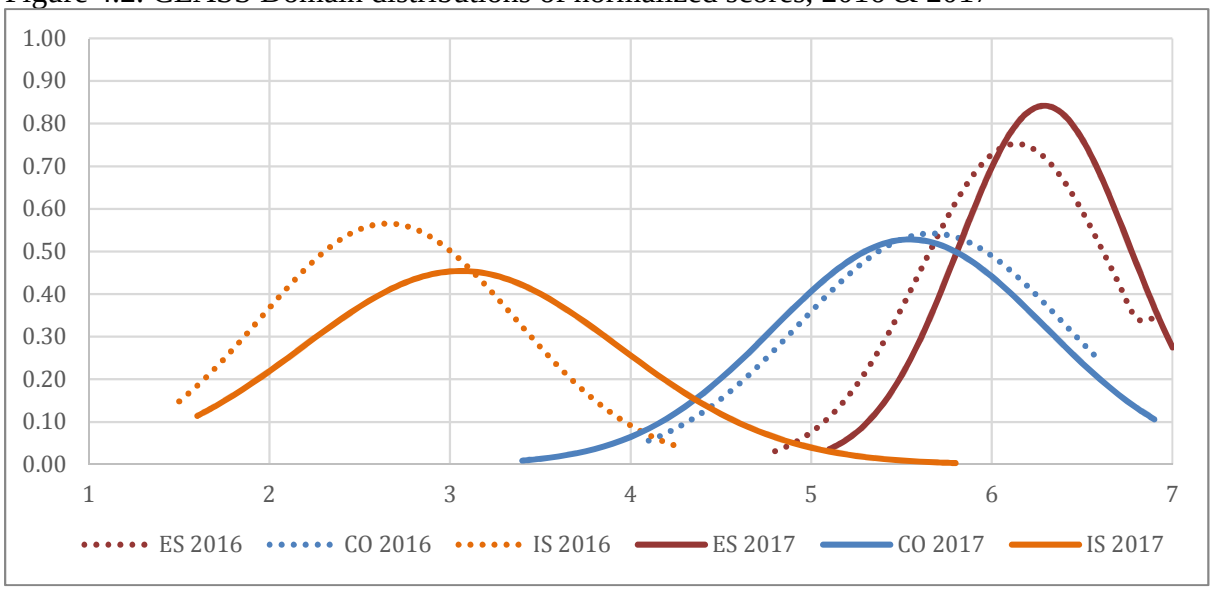


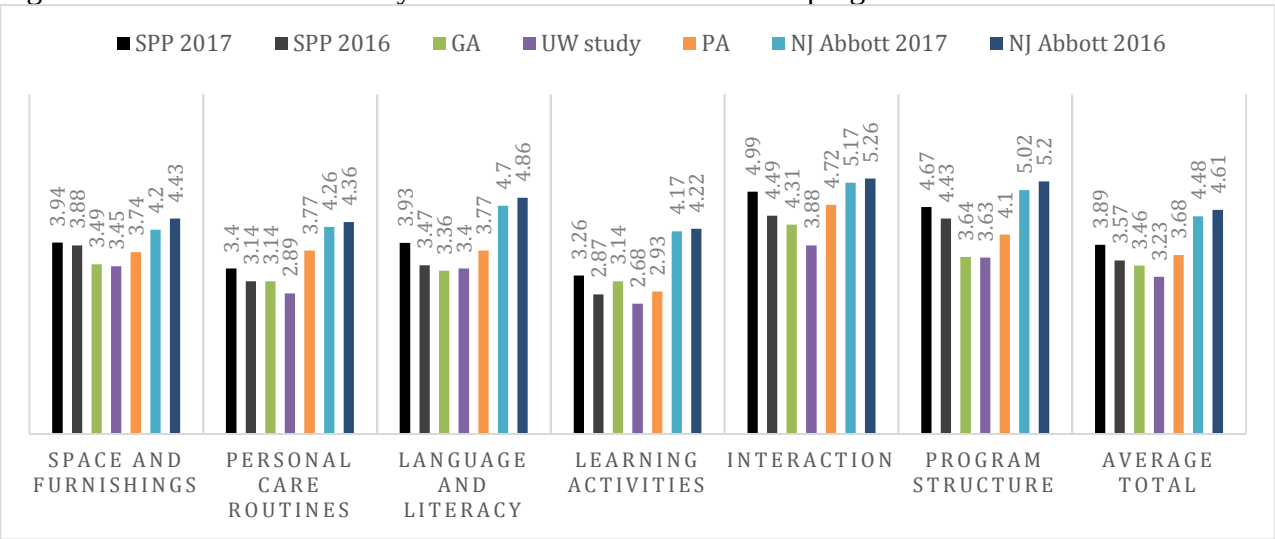
Table 8 and Figure 5 provide for context average ECERS-3 scores for 4 programs/studies: in GA, PA, UW state pre-K and childcare centers and NJ Abbott districts for this same year. In addition, for comparison, we have included ECERS-R data (which allows seeing growth over time for the two ECERS-R years provided) for two previous years for Abbott NJ districts. The ECERS-3 is still more the exception, rather than the rule in the field, with the ECERS-R still predominating, which does not allow for comparisons with many high-quality programs.

Table 8. Studies with reported ECERS-3 scores

Study	1. Space/ Furnishing	2. Personal Care Routines	3. Language & Literacy	4. Learning Activities	5. Interaction	6. Program Structure	Average Total
SPP 2017 (N=32)	**3.94 (0.61)**	**3.40 (0.86)**	**3.93 (0.82)**	**3.26 (0.57)**	**4.99 (1.07)**	**4.67 (0.86)**	**3.89 (0.55)**
SPP 2016 (N=12)	**3.88 (0.55)**	**3.14 (0.65)**	**3.47 (0.83)**	**2.87 (0.56)**	**4.49 (0.90)**	**4.43 (0.97)**	**3.57 (0.46)**
GA[1]	3.49	3.14	3.36	3.14	4.31	3.64	**3.46**
UW state pre-K & childcare (N=299)[2]	3.45	2.89	3.40	2.68	3.88	3.63	**3.23**
PA[3]	3.74	3.77	3.77	2.93	4.72	4.10	**3.68**
NJ Abbott:							
2016–17 (N=300)[4]	4.20 (0.84)	4.26 (1.14)	4.70 (1.10)	4.17 (1.11)	5.17 (1.30)	5.02 (1.38)	**4.48 (0.92)**
2015–16 (N=293)[3]	4.43 (1.02)	4.36 (1.33)	4.86 (1.26)	4.22 (1.17)	5.26 (1.34)	5.20 (1.31)	**4.61 (1.03)**
2007-08 (N=317)[5]	5.03	4.29	5.46	4.85	6.44	5.41	**5.20**
2002-03 (N=310)[4]	3.76	3.69	4.27	3.37	4.92	4.04	**3.96**

[1]Jenson (2015); [2]CQEL (Unpublished); [3]PAKEYS (Unpublished); [4]NIEER (2016); [5]NIEER (2017); [6]ECERS-R was used in these evaluations. Available at http://www.state.nj.us/education/ece/research/elichome.htm.

Figure 5. SPP ECERS scores by dimension in relation to other programs



We report CLASS scores for the SPP classrooms in 2016 and 2017, as well as for other programs in Table 9 and Figure 6 below. The 2017 SPP classroom average for emotional support is among the three top scoring ones (together with NYC and Pre-K 4 SA in San Antonio) and the scores for classroom organization are higher than several of these. On Instructional Support, SPP classrooms in 2017 still scored on average lower than most programs, but higher than the national average, and in line with the second year of the San Antonio program. The threshold suggested in the literature for quality (at 3) is lower in instructional supports than for other CLASS dimensions, and the SPP score is at the thresholds suggested in the literature (page 52).

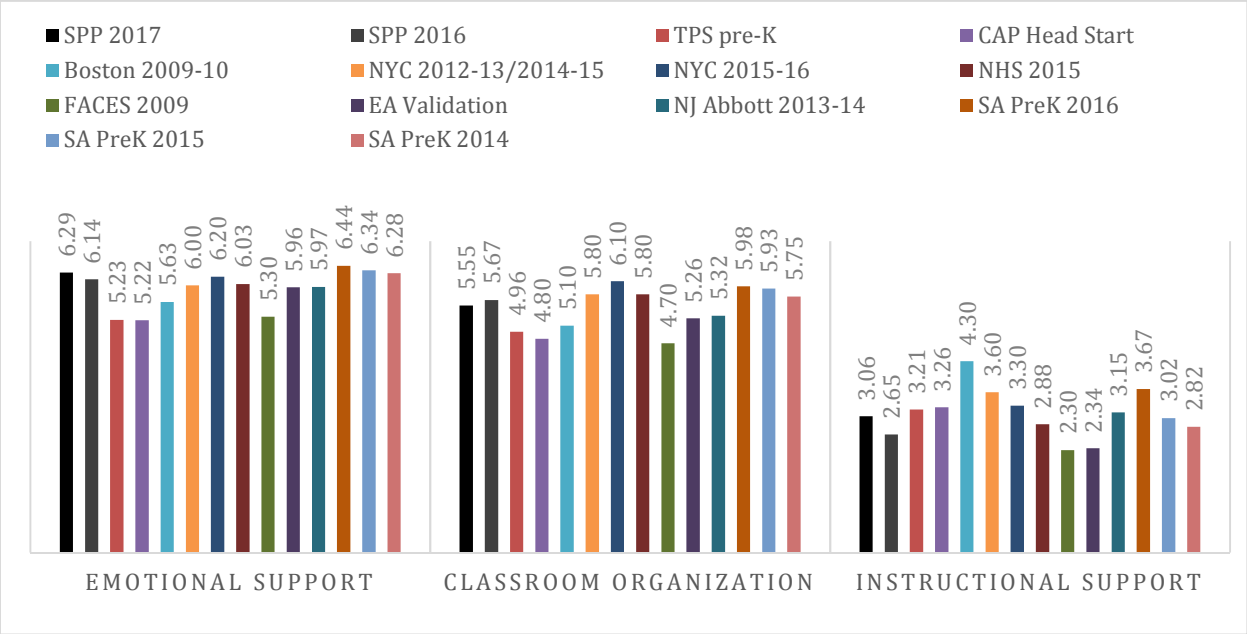
Table 9. Classroom quality across the nation, and for selected programs

Study	Emotional Support	Classroom Organization	Instructional Support
SPP classrooms 2017 (N=32)	**6.29 (0.47)**	**5.55 (0.76)**	**3.06 (0.88)**
SPP classrooms 2016 (N=14)	**6.14 (0.53)**	**5.67 (0.74)**	**2.65 (0.71)**
Tulsa[1]			
TPS pre-k (N=77)	5.23 (0.57)	4.96 (0.69)	3.21 (0.93)
CAP Head Start (N=28)	5.22 (0.78)	4.80 (0.84)	3.26 (0.94)
Boston[2] (N=83) (2009-2010)	5.63 (0.60)	5.10 (0.68)	4.30 (0.84)
NYC (N=555) (2012-13 to 2014-15)[3]	6.00	5.80	3.60
NYC (N=1,134) (2015–16)[3]	6.20	6.10	3.30
National Head Start Overview 2015[4]	6.03 (0.28)	5.80 (0.36)	2.88 (0.54)
Head Start FACES 2009[5]	5.30	4.70	2.30
EA Validation study (N=75)[6]	5.96 (0.66)	5.26 (0.77)	2.34 (0.71)
NJ Abbott 2013-2014 (N=163)[7]	5.97 (0.63)	5.32 (0.89)	3.15 (0.96)
San Antonio (N=89) (2016)[8]	6.44 (0.51)	5.98 (0.81)	3.67 (1.23)
San Antonio (N=76) (2015)[9]	6.34 (0.64)	5.93 (0.97)	3.02 (1.14)
San Antonio (N=36) (2014)[10]	6.28 (0.35)	5.75 (0.60)	2.82 (0.82)

[1]Phillips et. al (2009); [2]Weiland et. al (2013); [3]NYC Department of Education (2017); [4]Office of Head Start. (2015); [5] Aikens et. al (2013); [6]CQEL (Unpublished); [7]NIEER (2014); [8]EDVANCE (2016); [9]EDVANCE (2016); [10]EDVANCE (2014).

Figure 6 illustrates how the SPP program average compares to these other programs.

Figure 6. SPP CLASS scores by dimension in relation to other programs



ECERS-3 subscales

Table 10 presents items and subscales across 2016 and 2017, including the average score, and the range which illustrates the minimum and maximum scored by classrooms.

The *Space and Furnishings* subscale looks at the physical space of a classroom. Included are whether children have enough space and furniture, whether the arrangement of the furniture allows for learning and exploration and whether displays are meaningful and representative of the children in the class. Under the *space and furnishing* subscale, the "indoor space," "child-related display," and "space for gross motor play" where the items that evidenced lower scores this spring, while all other items increased scores. Four items under this subscale ranged starting at 1, indicating that some classrooms scored at the inadequate rating. Similarly, in four of them, there were classrooms that scored as excellent. Gross motor equipment increased slightly but remains the lowest scoring item in this scale. "Space for gross motor" and "gross motor equipment" have a time requirement of 15 minutes to receive credit in the "minimal" category of scoring and 30 minutes for "good."

The second subscale *Personal Care Routines*, addresses the health, hygiene and safety practices of the classroom. Under *personal care routines*, "meals/snacks" and "toileting and diapering" scores improved. However, all items evidenced classrooms scoring as inadequate, and the average score remained under the minimal range (3–5) for most of items and below it for "health practices." This item requires five specific times for hand washing including before and after using wet or shared sensory materials, and upon arriving in the classroom. The minimal scores on it reveal that hand washing procedures probably need more attention.

The *Language and Literacy* subscale addresses how staff direct activities and materials towards developing children's language and literacy skills. All items under this subscale increased in relation to the previous year. The lowest scoring item was "Becoming Familiar with Print" which averaged a 3.25. This item expects observing visible print being combined with pictures and staff taking dictation of children's words in a way that is interesting and engaging to children for the purpose of showing print as a useful tool. The "Staff Use of Books" item averaged at 3.50. To receive a score in the good (5) to excellent (7) range on this item all children are required to be actively engaged during story time.

The *Learning Activities* assesses the presence, variety, and accessibility of learning materials in the classroom for children, and simultaneously captures the extent to which teachers actively engage children with the different types of materials. Under this subscale, the average for "fine motor" and "art" where the highest, 4.47 and 4.28 respectively. However, these are still not reaching the level of "good" (5.00). This means that during an observation there was no evidence of children interacting with teachers using the respective materials for these items. For the other areas, the scores are even lower, with some areas such as "blocks," "nature/science," and "math" scoring under 3 (minimal). This means that while the quantity and quality of materials needed to score higher may be present (the item does not differentiate), without evidence of interactions during the three-hour observation period, the score cannot be higher than a 2.00 or 3.00. Under Learning Activities attention is paid to the way that the items define interest centers so that classrooms satisfy the material requirements. It also captures the extent to which teachers move in the classroom utilizing the materials to generate meaningful learning exchanges.

The *Interaction* subscale measures supervision of children during gross motor time, how teachers individualize teaching and learning and children and teachers' interactions. All items under this subscale improved except for "peer interaction," and all scored above a 4. The "staff-child" item scored in the good range, with no classrooms scoring as inadequate. Most of the items in this subscale are close to the good range and working towards providing children with opportunities toward more selection of peers, resolving conflict, teachers explaining expectations for behavior, and responding to discipline issues explaining them and with care are aspects that would bring a couple of the items in this scale above the good range.

The last subscale is *Program Structure* which is centered on the general formats of the classroom and how the children spend their time. Only the item "whole-group activities for play and learning" increased under this scale, and significantly so, averaging now 4.81.

Table 10. ECERS-3 Item, Subscale, and Overall Means and Ranges by Item, 2016 & 2017

ECERS-3 Item and Subscales	2016 Mean (Range) N=14	2017 Mean (Range) N=32
Space and Furnishings		
1. Indoor space	6.43 (4-7)	5.47 (2-7)
2. Furnishings for care, play and learning	4.36 (4-7)	4.56 (3-7)
3. Room arrangement for play and learning	3.64 (2-7)	4.72 (2-7)
4. Space for privacy	4.14 (2-6)	4.53 (1-7)
5. Child-related display	3.36 (1-5)	3.09 (1-4)
6. Space for gross motor play	3.14 (1-4)	3.06 (1-6)
7. Gross motor equipment	2.07 (1-4)	2.13 (1-5)
Personal Care Routines		
8. Meals/ snacks	3.07 (1-4)	3.88 (1-7)
9. Toileting/diapering	2.21 (1-3)	3.19 (1-7)
10. Health practices	2.93 (2-4)	2.69 (1-5)
11. Safety practices	4.36 (2-7)	3.88 (1-7)
Language and Literacy		
12. Helping children expand vocabulary	3.50 (3-5)	3.63 (1-7)
13. Encouraging children to use language	4.36 (3-7)	4.84 (3-7)
14. Staff use of books with children	3.07 (1-6)	3.50 (1-6)
15. Encouraging children's use of books	4.21 (1-7)	4.41 (3-6)
16. Becoming familiar with print	2.21 (1-4)	3.25 (1-6)
Learning Activities		
17. Fine motor	4.36 (2-5)	4.47 (2-7)
18. Art	3.71 (2-6)	4.28 (1-7)
19. Music and movement	3.50 (2-5)	3.47 (2-6)
20. Blocks	2.00 (1-4)	2.97 (1-5)
21. Dramatic Play	2.79 (1-6)	3.50 (1-7)
22. Nature/science	2.50 (1-4)	2.28 (1-5)
23. Math materials and activities	1.71 (1-3)	2.25 (1-4)
24. Math in daily events	2.86 (1-5)	3.34 (1-5)
25. Understanding written numbers	1.29 (1-2)	1.69 (1-5)
26. Promoting acceptance of diversity	4.21 (3-6)	4.34 (2-6)
Interaction		
27. Appropriate use of technology	N/A (1-1)*	N/A
28. Supervision of gross motor	3.71 (1-7)	4.56 (1-7)
29. Individualized teaching and learning	4.21 (3-7)	4.94 (2-7)
30. Staff-child interaction	4.93 (3-7)	5.66 (3-7)
31. Peer interaction	5.00 (3-7)	4.84 (1-7)
32. Discipline	4.57 (2-7)	4.97 (2-7)
Program Structure		
33. Transitions and waiting times	4.86 (3-7)	4.75 (3-7)
34. Free play	4.50 (3-6)	4.44 (2-7)
35. Whole - group activities for play and learning	3.93 (2-5)	4.81 (2-6)

Note: (*) Only 2 classrooms received a score for #27, both were 1. All others were N/A.

CLASS: Emotional Support Domain

Table 11 shows the score dimensions under the three CLASS domains. The Emotional Support (ES) domain focuses on how the teacher fosters a nurturing and safe environment for children to learn. The "Positive Climate" and "Negative Climate" dimensions assess the emotional connection between teachers and students. Positive Climate "reflects the emotional connection between the teacher and students and among students and the warmth, respect, and enjoyment communicated by verbal and nonverbal interactions" (Pianta, La Paro & Hamre, p.23). Negative Climate "reflects the overall level of expressed negativity in the classroom" (p. 28). Negative Climate scores have been inverted throughout this report, and scores the highest, indicating a lack of expressed negativity. "Positive Climate" increased relative to the previous year (average 6.33) and scores were closely aligned with the negative climate domain.

The "Teacher Sensitivity" dimension captures whether teachers are able to anticipate problems and provide support for children (average 6.04). The average score for this dimension is now in the high range, evidencing with slightly more consistency teachers mostly aware of children, responsive to their needs and emotions, providing individualized support, addressing problems and generally comforting children.

"Regard for Student Perspectives" (average 5.96, just like the previous year) focuses on how comfortable students appear to be in their classroom environment. This is assessed based on how children participate, seek help and take risks, on whether teachers foster an environment where children feel safe to do all these things, on the degree to which interactions are based on children's interests and perspectives, and how well teachers encourage child autonomy. More consistent opportunities for children to have time to express themselves and move about freely in the classroom, encouraged by the teacher, and with the teacher developing interactions based on child interests would bring this score into the higher range.

CLASS: Classroom Organization Domain

The Classroom Organization domain assesses the supports through which the teachers manage behavior and redirect it, manage instructional time and routines, and manage activities and takes advantage of students' interests. "Behavior Management" focuses on whether behavior expectations are clear and consistent, and on how proactive teachers are in preventing misbehavior. "Productivity" measures teachers' time management, pacing and transitions throughout the day and across activities. "Instructional Learning Formats" measures how teachers maximize their facilitation of student learning during activities. The latter includes how effective questions are, clear learning objectives, and the range of opportunities for children to learn. Student interest is also accounted for here.

This domain had a slight decrease relative to the previous year. "Productivity" scored the highest (average 5.91), and no classrooms scored under 3 for any of the dimensions. A score in the mid-range signifies that while the majority of the time children are provided with activities, there are periods of time lost with disruptions or in transitions. The lowest scoring dimension is "Instructional Learning Formats" (average 5.21, the same as the previous year), also in the mid-range score. Increasing this dimension would require more consistent active engagement and facilitation from the teacher using varied materials and modalities. Also, the teacher should

effectively be focusing students on learning objectives and students should be seen consistently engaged or interested in the activities without this interest waning.

CLASS: Instructional Supports Domain

The Instructional Supports Domain measures the interactions through which teachers deliver and facilitate high-order thinking skills, provide feedback, encourage participation and develop language. This domain is the most difficult, yet most important, when considering teacher practices that bare impacts on student growth. The complexity involved in this domain leads to low scores in it consistently across most programs. All three dimensions under this domain increased relative to the previous year.

The dimension of "Concept Development" measures teachers' use of discussions to stimulate reasoning and analysis and encourage understanding. It also inquires into teachers' encouragement of creativity and integration of concepts into children's lives. Like the other dimensions in CLASS, it is key for a high score that teachers are consistent and intentional, rather than thing occurring in isolation or sometimes. Concept Development scored the lowest (average 2.64). Supporting this dimension would require more frequent and much more consistent use of discussions and activities that foster reasoning and analysis by children, opportunities for children to create and generate their own ideas (or products) and teachers relating new concepts to those previously learned or students' lives.

"Quality of Feedback" measures the quality of teacher responses to children's talk, that is, whether teachers provide hints, are persistent, ask for explanations of thinking, and how specific they are in responses to children. Classrooms scored just about in the mid-range (average 3.03). Supporting increases in quality would require for teachers to scaffold children, helping them solve problems by providing resources or asking questions, and doing up until the child comes to a solution. It also requires that teachers consistently scaffold children through their problem-solving process or through understanding a concept, that they engage them in feedback loops regularly, and teacher expand on what children say, prompting them to explain their own thinking.

"Language Modeling" is the last dimension of the CLASS and measures both the quality and quantity of teacher's language used to develop language in children (average 3.57). Mid-high range classrooms on this dimension demonstrate frequent conversations between teachers and children, many open-ended questions, and the use of self and parallel talk when working with children in play areas, the use of advanced language with students including the use of varied words and the introduction of new words.

Table 11. CLASS Dimension and Domain Means and Range by Item, 2016 & 2017

CLASS Dimensions and Domains	2016 Mean (Range) N=14	2017 Mean (Range) N=32
Emotional Support Domain	6.14 (4.88-6.81)	6.29 (5.19-7.00)
1. Positive Climate	5.80 (4.25-7.00)	6.33 (5.25-7.00)
2. Negative Climate*	6.86 (5.75-7.00)	6.95 (6.63-7.00)
3. Teacher Sensitivity	5.91 (4.25-6.75)	6.04 (4.25-7.00)
4. Regard for Student Perspectives	5.96 (4.25-7.00)	5.96 (4.25-7.00)
Classroom Organization Domain	5.67 (4.17-6.58)	5.55 (3.42-6.83)
5. Behavior Management	5.73 (3.75-7.00)	5.46 (3.50-6.75)
6. Productivity	6.05 (4.50-7.00)	5.91 (3.50-7.00)
7. Instructional Learning Formats	5.21 (3.50-6.50)	5.21 (3.00-6.75)
Instructional Support Domain	2.65 (1.50-4.25)	3.06 (1.67-5.75)
8. Concept Development	2.07 (1.25-3.50)	2.64 (1.25-5.50)
9. Quality of Feedback	2.61 (1.50-4.25)	3.03 (1.50-5.50)
10. Language Modeling	3.29 (1.75-5.00)	3.57 (1.75-6.25)

Note: (*) The Negative Climate dimension was transposed so that on here, high represents “good”.

3. How does quality vary within SPP and do children from different backgrounds experience different quality?

Figures 7 and 8 illustrate average classroom quality scores for ECERS-3 and all three CLASS domains across agencies. For the most part, score patterns are quite similar, with ECERS scores between 3 and 4 for most agencies, except for two agencies scoring above 4 and one above 5. CLASS ES & CO scores in the 5–7 range across all agencies (one agency, coded as ten, slightly lower than the rest). CLASS IS scores do appear to vary across agencies, with four in the 2-3 range, four in the 3–4 range and two at or above 4. Scores by Agency are reported in Appendix A, Tables A.1 and A.2 (which also include the previous year). Tests of statistical significance between groups found no differences across agencies on the ECERS-3, the CLASS ES and the CLASS IS. For the CLASS CO, there were significant differences between Agencies 2, 3, 4 and 10 as these scored higher or lower than average CLASS CO scores.

Figure 7. ECERS scores by Agency

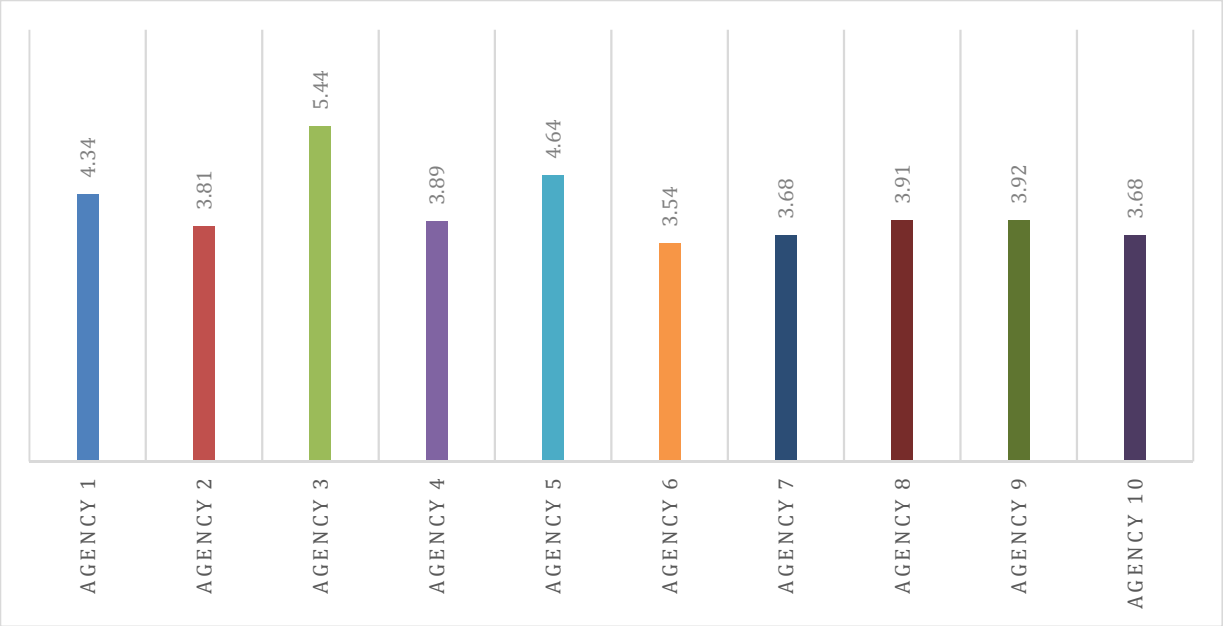
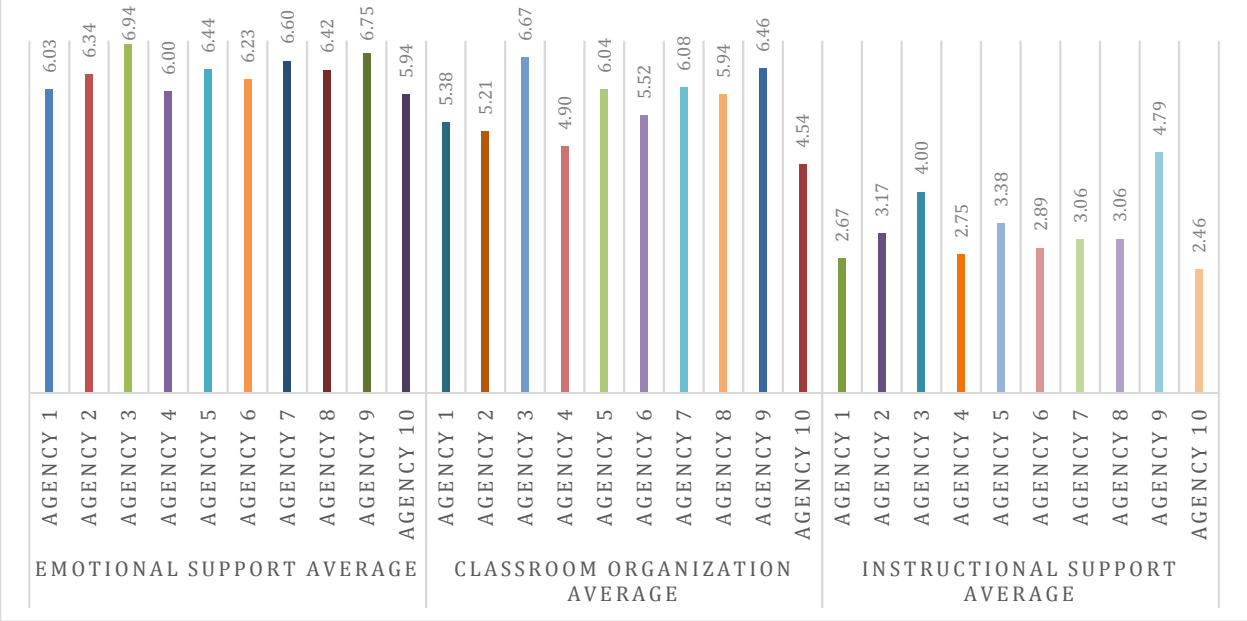


Figure 8. CLASS domain scores by Agency



Similarly, Figure 9 illustrates ECERS-3 and CLASS domain scores for smaller (classrooms with 18 or less children) and larger (with more than 18 children) classrooms in the sample (scores are reported in Appendix Tables C.3 and C.4). Overall, classroom quality patterns are very close together regardless of class size, between 3 and 4 for ECERS, about 6 for CLASS ES, and between 5 and 6 for CLASS CO. There is a slightly larger difference for CLASS IS with classrooms under 18 children scoring on average 3.13 and classrooms above 18 children scoring

on average 2.85. Tests of statistical significance between groups found no differences in scores between smaller and larger classrooms.

Figure 9. ECERS and CLASS Domain scores by Class Size

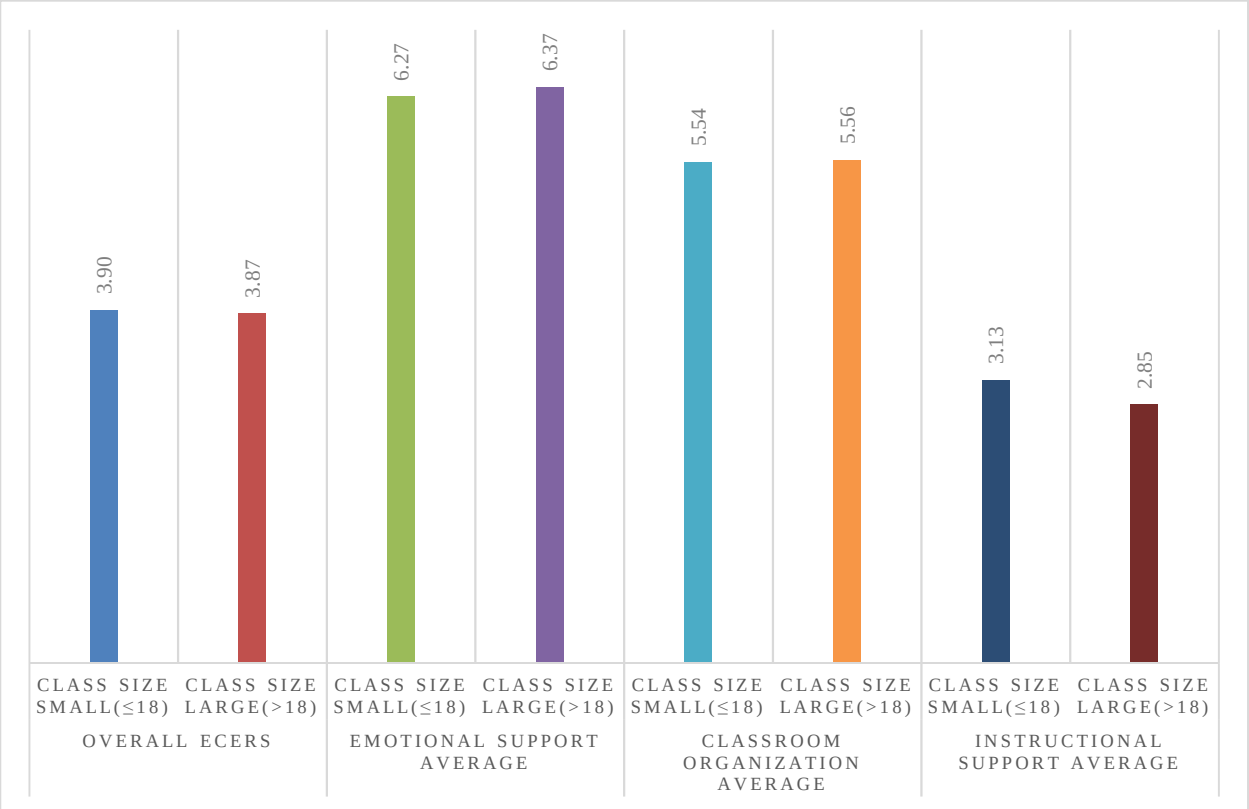


Figure 10 illustrates the quality of care by children's gender, ethnicity/race, language background and FPL for the SPP children in the sample. We observe no distinct patterns by child characteristics with any one group exposed to better/lower quality than their peers for ECERS-3, CLASS ES and CO. All children seem to be receiving equivalent levels of quality of care across these. Tests of statistical significance between groups found no significant differences in quality by gender, ethnicity or language. The only statistically significant difference found was for income and CLASS CO, with the families under 100% FPL and between 100–300% FPL attending classrooms with statistically lower CLASS CO levels than families above 300% FPL.

Figure 10. ECERS and CLASS Domain scores by Child Characteristics

Classroom quality by year of entry into SPP

We inquired into whether there were differences in quality between new classrooms in the program, and those with a year in the program. Tables 12 and 13 describe ECERS-3 and CLASS scores for classrooms grouped according to the number of years in SPP. Classrooms with 2 years in the program scored slightly higher in the ECERS-3 and reach higher scores than newer classrooms, while this is inverted for CLASS domains where newer classrooms were the ones scoring slightly lower. Test of statistical significance between groups found no differences between ECERS-3 and CLASS domains across these two groups of classrooms.

Table 12. ECERS-3 Subscale, and Overall Means and Ranges, 2017 (N=32)

ECERS-3 Item and Subscales	2 years in SPP (N=9)				1 year in SPP (N=23)			
	Mean	(SD)	Min.	Max.	Mean	(SD)	Min.	Max.
Overall	**3.93**	**0.63**	**3.32**	**5.44**	**3.87**	**0.53**	**2.74**	**5.09**
Space and Furnishings	4.06	0.58	3.43	5.29	3.89	0.63	2.71	5.14
Personal Care Routines	3.89	0.76	2.75	5.50	3.22	0.83	1.50	4.50
Language and Literacy	4.13	0.90	2.80	6.00	3.84	0.79	2.40	5.60
Learning Activities	3.10	0.61	2.40	4.50	3.32	0.55	2.60	4.70
Interaction	4.96	1.09	3.60	6.80	5.01	1.09	2.40	6.80
Program Structure	4.59	0.94	3.67	6.33	4.70	0.87	3.00	6.00

Table 13. CLASS Domain Means and Ranges, 2017 (N=32)

CLASS Domains	2 years in SPP (N=9)				1 year in SPP (N=23)			
	Mean	(SD)	Min.	Max.	Mean	(SD)	Min.	Max.
Emotional Support	6.28	0.47	5.31	6.94	6.29	0.49	5.19	7.00
Classroom Organization	5.43	0.71	4.42	6.67	5.59	0.78	3.42	6.83
Instructional Support	3.00	0.58	2.42	4.00	3.08	0.98	1.67	5.75

Classroom quality for children in the control group

Classroom quality for children in the control group is shown for the ECERS-3 and its subscales (Table 14) and for the CLASS domains (Table 15). The quality experienced by children in the control group recruited for this study (where children in the SPP waiting list attended AY 2016–17) was lower for both ECERS-3 and all three CLASS domains, and for all but one subscale in the ECERS-3 (program structure), relative to SPP quality. In addition, SPP had higher maximum scores on the ECERS-3 (and all its subscales) as well as on CLASS ES and CLASS IS.

Table 14. ECERS-3 Subscale, and Overall Means and Ranges, 2017 (N=7)

Variable	Mean	(SD)	Min.	Max.
Overall	**3.51**	**0.75**	**2.29**	**4.32**
Space and Furnishings	3.27	1.41	1.43	4.86
Personal Care Routines	3.18	1.00	1.50	4.50
Language and Literacy	3.91	0.73	2.60	5.00
Learning Activities	2.50	0.57	1.90	3.60
Interaction	4.97	1.11	3.60	6.40
Program Structure	4.81	0.77	4.00	6.00

Table 15. CLASS Domain Means and Ranges, 2017 (N=7)

CLASS Domains	Mean	(SD)	Min.	Max.
Emotional Support	6.21	0.61	5.06	6.94
Classroom Organization	5.49	1.12	3.50	6.83
Instructional Support	2.40	0.83	1.25	3.50

4. How did the learning of children enrolled in SPP classrooms progress in 2016–17, and how did it vary with classroom quality?

This evaluation reports standardized measures of child outcomes in two content areas: receptive vocabulary (using the Peabody Picture Vocabulary Test) and literacy (using the Woodcock-Johnson Tests of Achievement Letter-Word subtest), as well as math (using the Woodcock-Johnson Tests of Achievement Applied Problems subtest). In addition, it reports on two measures of executive functioning (EF): Dimensional Change Card Sort Game (DCCS) and Peg Tapping task (PT). The latter two assess a combination of short-term memory, the ability to inhibit automatic response tendencies that can interfere with achieving a task, and the capacity for set shifting.

We organize descriptive results from the 2016–17 evaluation by first showing gains for the SPP sample and then split out by various child subgroups, by agency, comparing classrooms with class sizes under 18 with classrooms with class sizes above 18, and comparing classrooms below a threshold for high quality. The statistical significance for these groups is measured further below through estimations that relate these characteristics to children's gains in the various measures included in the study. We also compared these to children's gains in 2015–16, with the caveat that in 2015–16 all children were assessed while in 2016–17 only a random sample of children was assessed. Children's learning gains are set in contrast to average gains reported in FACES (Aikens, Klein, Tarullo, & West, 2013). Finally, we report results from multivariate analyses that examines variations in outcomes with all the child and program characteristics simultaneously, and captures the differences between SPP and non-SPP children in the program. Children from Spanish speaking homes were tested in Spanish as well as in English and estimations using their Spanish language vocabulary did not change any of the results presented in this report. Receptive vocabulary measured by the PPVT is presented first, followed by literacy (WJ-LW), early math (WJ-AP), and two measures of executive functioning, the Dimensional Change Card Sort Game (DCCS) and the Peg Tapping task (PT).

Receptive vocabulary results

Table 16 shows children's vocabulary scores results for the fall (pre-test) and spring (post-test) and the gains from fall to spring. Standardized scores—which are adjusted for age—are reported in this section (raw scores are reported in Appendix B.1, Table B.1.1). The mean standard score for this measure is set at 100 which is another way of saying that the average child in the U.S. population is expected to score 100 at any age. The standard deviation is 15. Thus, positive gains are an indication that children improved more over the course of the preschool year than is expected based on the change in age alone. Information on this table reflects the performance of all children regardless of language background. We only report scores for children with valid scores in both the fall and spring of the school year.

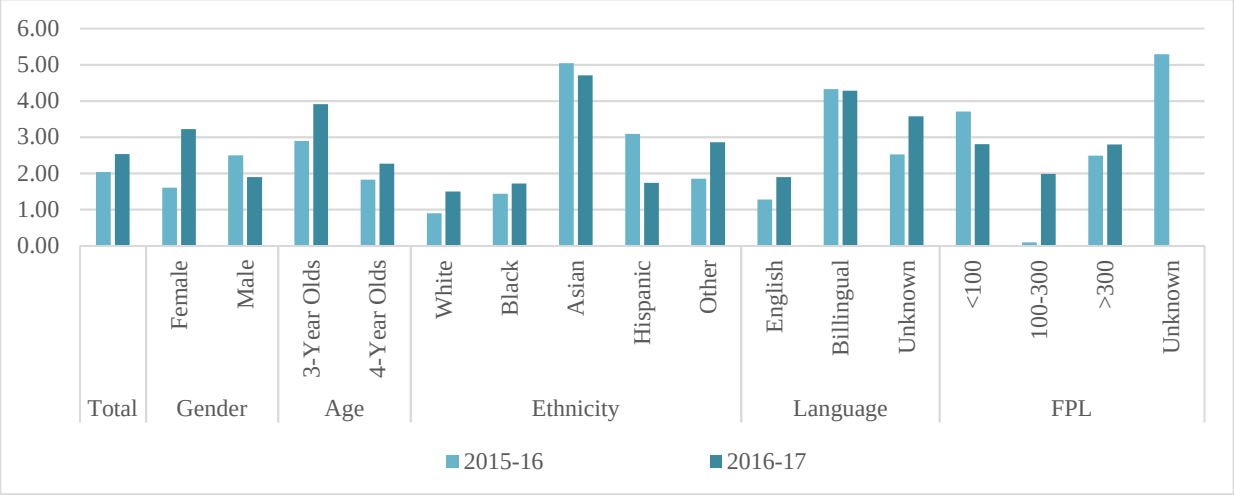
On the whole, children scored at the population average in the fall and slightly above the average in the spring. One-year gains for the SPP sample were of 2.53 standard points; slightly more than half of the 4.5 point one-year gains reported for 4-year-olds in the FACES study, although Head Start children scored well below average before and after a year in the program (Table B.5a; Aikens, Klein, Tarullo, & West, 2013). Minority children score considerably below average and make larger gains, particularly Asian children and children of mixed or other backgrounds. Children speaking languages other than English score the lowest and make the largest gains. Children below the poverty level start lower than their higher income peers but also make the largest gains, ending up near the national average. The statistical significance for these differences is assessed on the multivariate analyses on pages 47-49. For comparison, FACES reported larger 2009 PPVT-4 standard gains for four-year-olds, with 3.4 points for White children, and 4.3 points for Black children (see Barnett, 2013). However, children in FACES start at a much lower level and even with the larger gains do not approach the national average.

Table 16. Receptive vocabulary means and gains by child characteristics

		N	PPVT 2016 Fall		PPVT 2017 Spring		PPVT Gains 2016–17	
			Mean	SD	Mean	SD	Mean	SD
Total		288	100.68	17.98	103.21	16.85	2.53	9.57
Gender	Female	139	101.71	17.51	104.93	15.86	3.22	9.16
	Male	149	99.72	18.41	101.6	17.62	1.89	9.92
Age	3-Year-Old Cohort	46	92.22	15.36	96.13	13.66	3.91	9.91
	4-Year-Old Cohort	242	102.29	18.01	104.55	17.08	2.26	9.50
Ethnicity	White	61	116.56	13.69	118.05	11.97	1.49	9.41
	Black	66	91.38	14.99	93.09	14.42	1.71	9.72
	Asian	48	92.60	17.79	97.31	16.53	4.71	8.61
	Hispanic	22	89.59	17.54	91.32	14.09	1.73	9.61
	Other	91	103.73	14.19	106.58	13.47	2.86	10.03
Language	English	198	105.81	16.53	107.7	16.17	1.89	10.07
	DLLs	46	88.20	14.89	92.48	14.56	4.28	6.79
	Unknown	44	90.66	16.81	94.23	13.27	3.57	9.61
FPL	<100	59	91.44	15.76	94.24	15.47	2.80	7.99
	100-300	93	108.95	18.20	110.92	17.11	1.98	9.50
	>300	136	99.04	16.36	101.82	14.93	2.79	10.27

Figure 11 compares gains for children in the 2016–17 sample, with gains for children enrolled in SPP in 2015–16 by subgroup. Standard gains were 0.50 standard points larger this year, driven by larger increases for females and 3-year-olds. The large decrease in gains for Hispanics may be the consequence of a change in the size of this group relative to last year. Note that comparison using race/ethnicity based on DEEL demographics (Appendix B.2, Table B.2.1) do not differ much in fact.

Figure 11. Receptive vocabulary gains by child characteristics and year



Children's pre-test and post-test vocabulary standard scores for selected center characteristics are reported in Table 17 (raw scores are reported in Appendix B.1, Table B.1.2). Children in higher quality classrooms, as measured by the observational measures, evidence higher average gain patterns in CLASS CO and IS domains. Very few children were in

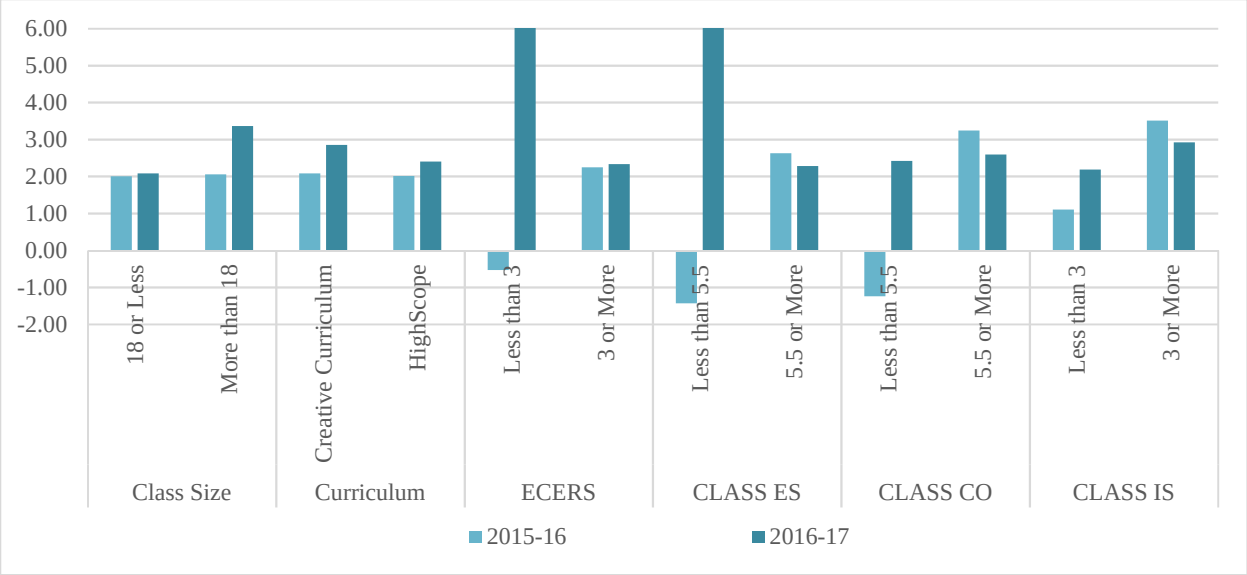
classrooms with ECERS-3 scores under 3, and CLASS ES scores under 5.5, although these children did experience high gains regardless.

Table 17. Receptive vocabulary means and gains by center characteristics

		N	PPVT 2016 Fall		PPVT 2017 Spring		PPVT Gains 2016–17	
			Mean	SD	Mean	SD	Mean	SD
Total		288	100.68	17.98	103.21	16.85	2.53	9.57
Class Size	18 or Less	187	99.45	18.30	101.53	17.06	2.08	9.79
	More than 18	101	102.95	17.22	106.31	16.06	3.36	9.13
Curriculum	Creative Curr.	84	99.06	20.06	101.90	19.10	2.85	10.03
	HighScope	204	101.35	17.05	103.75	15.84	2.40	9.40
ECERS	Less than 3	14	96.00	16.98	102.43	16.94	6.43	8.67
	3 or More	274	100.92	18.02	103.25	16.87	2.33	9.59
CLASS ES	Less than 5.5	16	97.63	17.05	104.44	16.36	6.81	9.78
	5.5 or More	272	100.86	18.04	103.14	16.90	2.28	9.52
CLASS CO	Less than 5.5	106	98.51	16.85	100.92	17.10	2.42	9.95
	5.5 or More	182	101.95	18.53	104.54	16.60	2.59	9.37
CLASS IS	Less than 3	156	99.34	18.41	101.53	17.06	2.19	10.32
	3 or More	132	102.27	17.38	105.19	16.43	2.92	8.62

Figure 12 illustrates gains for these same classroom characteristics by year. The critical improvements here are important increases in gains standard PPVT scores in lower ECERS and CLASS quality classrooms. Higher quality classrooms on the CLASS however do not follow this pattern.

Figure 12. Receptive vocabulary gains by center characteristics and year



Literacy results

Children's WJ-III letter-word (LW) identification scores for the overall sample and by selected child characteristics are reported in Table 18. The LW subtest measures children's ability to

identify letters and subsequently read a list of words of increasing difficulty. The test also has a mean standard (i.e., age adjusted score) of 100 and a standard deviation of 15 (raw scores are reported in Appendix B.1, Table B.1.3). Scores for all children with a valid fall and spring score are included regardless of their home language background.

Children in the SPP sample scored on average slightly above the norms in both the fall and the spring. One-year gains for the whole group of children were of 1.07 standard points. This is equivalent to a fifth of the reported one-year gains for 4-year-olds in the FACES study of 5.0 standard points, with Head Start children in FACES scoring below the average of 100 (Table B.5a; Aikens et. al, 2013). In terms of specific groups, 3-year-olds, Black, dual-language children, and children under the FPL evidenced larger gains. For comparison, FACES reported for Head Start in 2009 LW standard gains for four-year-olds of 4.3 for White children, and 4.8 for Black children (see Barnett, 2013).

Table 18. Literacy means and gains by child characteristics

		N	WJ-LW 2016 Fall		WJ-LW 2017 Spring		WJ-LW Gains 2016–17	
			Mean	SD	Mean	SD	Mean	SD
Total		286	100.87	15.90	101.94	15.40	1.07	10.06
Gender	Female	138	100.37	15.48	101.85	15.54	1.48	10.23
	Male	148	101.34	16.33	102.03	15.32	0.70	9.93
Age	3-Year-Old Cohort	47	99.38	14.86	102.62	13.04	3.23	12.68
	4-Year-Old Cohort	239	101.16	16.11	101.81	15.84	0.65	9.44
Ethnicity	White	61	105.98	15.47	106.31	14.78	0.33	7.88
	Black	65	98.62	16.48	101.69	16.50	3.08	11.28
	Asian	48	104.44	16.55	106.23	16.92	1.79	11.51
	Hispanic	21	89.76	13.22	91.33	10.68	1.57	11.39
	Other	91	99.74	14.42	99.38	13.46	-0.35	9.23
Language	English	196	101.38	15.81	101.71	15.24	0.33	9.34
	DLLs	46	102.91	16.50	105.37	17.54	2.46	9.28
	Unknown	44	96.45	15.22	99.41	13.31	2.95	13.33
FPL	<100	59	97.41	16.07	101.61	13.63	4.20	12.47
	100-300	93	102.51	16.25	102.22	15.54	-0.29	7.97
	>300	134	101.26	15.47	101.90	16.13	0.64	9.97

Figure 13 displays these gains in standard WJ-LW scores in relation to those of children enrolled in SPP the year before. Average gains this year were 1.5 standard points lower. Some specific groups had increased gains in relation to the previous years, such as Blacks, Asians, Bilingual Children and Children under 100% FPL.

Figure 13. Literacy by child characteristics and year

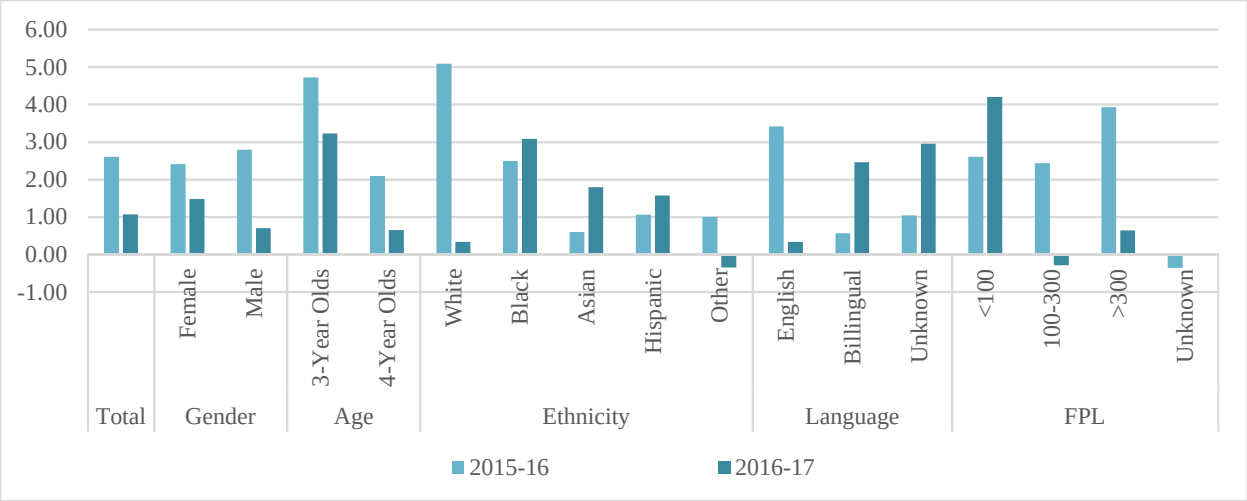


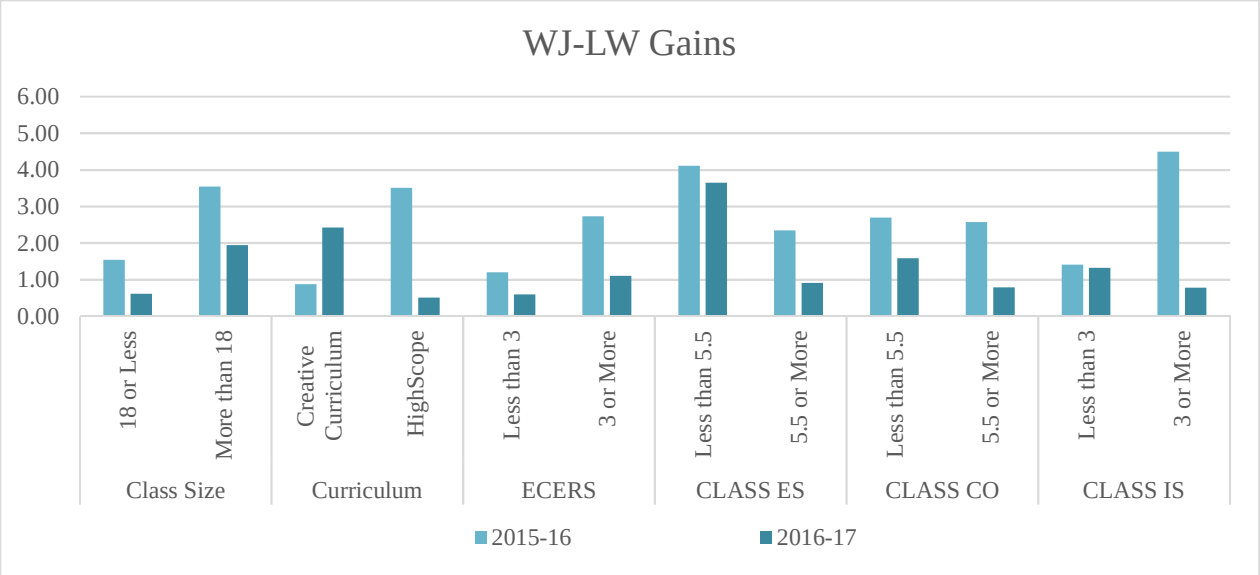
Table 19 reports SPP children's pre- and post-test letter-word identification standard scores across selected center characteristics (raw scores are reported in Appendix B.1, Table B.1.4). Children's gains differ across agencies and classroom characteristics and are higher when ECERS is above 3, and in classrooms implementing Creative Curriculum.

Table 19. Literacy means and gains by center characteristics

		N	WJ-LW 2016 Fall		WJ-LW 2017 Spring		WJ-LW Gains 2016–17	
			Mean	SD	Mean	SD	Mean	SD
Total		288	100.87	15.90	101.94	15.40	1.07	10.06
Class Size	18 or Less	187	101.03	16.44	101.65	15.39	0.61	10.89
	More than 18	99	100.57	14.91	102.51	15.49	1.94	8.27
Curriculum	Creative Curr.	84	100.68	17.4	103.1	16.69	2.42	12.08
	HighScope	102	100.95	15.28	101.47	14.85	0.51	9.07
ECERS	Less than 3	15	98.00	15.07	98.60	8.69	0.60	10.18
	3 or More	271	101.03	15.96	102.13	15.68	1.10	10.08
CLASS ES	Less than 5.5	17	96.82	13.81	100.47	12.87	3.65	10.31
	5.5 or More	269	101.13	16.01	102.04	15.56	0.91	10.05
CLASS CO	Less than 5.5	103	98.27	14.76	99.85	13.47	1.58	11.07
	5.5 or More	183	102.33	16.37	103.12	16.31	0.79	9.47
CLASS IS	Less than 3	154	100.80	15.53	102.12	14.79	1.32	10.19
	3 or More	132	100.95	16.39	101.73	16.14	0.78	9.95

These gains are shown in comparison to those of the previous year in Figure 14. Overall, most types of classrooms whether seen by size, or quality, children in the sample evidenced lower average child gains in literacy this year relative to the previous one.

Figure 14. Literacy gains by center characteristics and child



Early math results

We report children's pre- and post-test math scores, as measured by the applied problems (AP) subscale of the WJ-III in Table 20. Like the two measures above, AP is normed with a mean of 100 and a standard deviation of 15. On average, children in the SPP sample scored above average in the fall and spring of the school year. One-year gains for the whole group of children were of 2.21 standard points (average raw score gains are reported in Table B.1.5 in appendix B.1). This equates FACES one-year gains for 4-year-olds of 2.2 standard points (although Head Start children in such study scored below the norm throughout; Table B.5a; Aikens et. al, 2013). Among children in the sample, 3-year-olds, minorities, and dual language learners outperformed their peers in SPP. Gains for all subgroups of children differ strongly with those reported in FACES, where in 2009 AP standard gains for 4-year-olds were 1.4 points for White children, and 0.6 points for Black children (see Barnett, 2013). Negative "gains" for White children indicate that they lost ground relative to expectations for their higher age at post-test. Appendix B.1, Table B.1.5 shows raw scores gains are positive.

Table 20. Math means and gains by child characteristics

		N	WJ-AP 2016 Fall		WJ-AP 2017 Spring		WJ-AP Gains 2016–17	
			Mean	SD	Mean	SD	Mean	SD
Total		286	102.37	14.90	104.58	12.89	2.21	12.28
Gender	Female	138	102.01	15.71	103.43	12.59	1.42	12.29
	Male	148	102.70	14.15	105.66	13.12	2.95	12.27
Age	3-Year-Old Cohort	47	94.02	15.37	101.74	12.92	7.72	15.98
	4-Year-Old Cohort	239	104.01	14.27	105.14	12.84	1.13	11.14
Ethnicity	White	61	113.10	11.71	110.30	10.94	-2.82	12.20
	Black	65	94.00	14.39	98.62	11.82	4.62	13.12
	Asian	48	100.73	15.81	104.94	14.59	4.21	13.56
	Hispanic	21	92.19	13.22	97.38	11.24	5.19	8.70
	Other	91	104.36	11.70	106.49	12.00	2.13	10.84
Language	English	196	105.71	13.10	106.41	12.66	0.70	11.83
	DLLs	46	97.54	16.61	102.87	14.44	5.33	11.11
	Unknown	44	92.52	15.15	98.25	9.86	5.70	14.25
FPL	<100	59	95.08	14.39	98.88	10.54	3.78	12.56
	100-300	93	107.80	13.30	108.37	13.26	0.57	11.35
	>300	134	101.81	14.79	104.47	12.71	2.66	12.73

The figure below illustrates gains for each of these subpopulation groups in the sample, in relation to the previous year. The most remarkable difference is that for the overall sample, as well as for every group that in 2015–16 there was an observed negative standard gain, this year, this was reversed. The only exception being White children in the sample. This made the difference in gains between last year and this year of 3.75 standard points for the WJ-AP.

Figure 15. Math gains by child characteristics and year

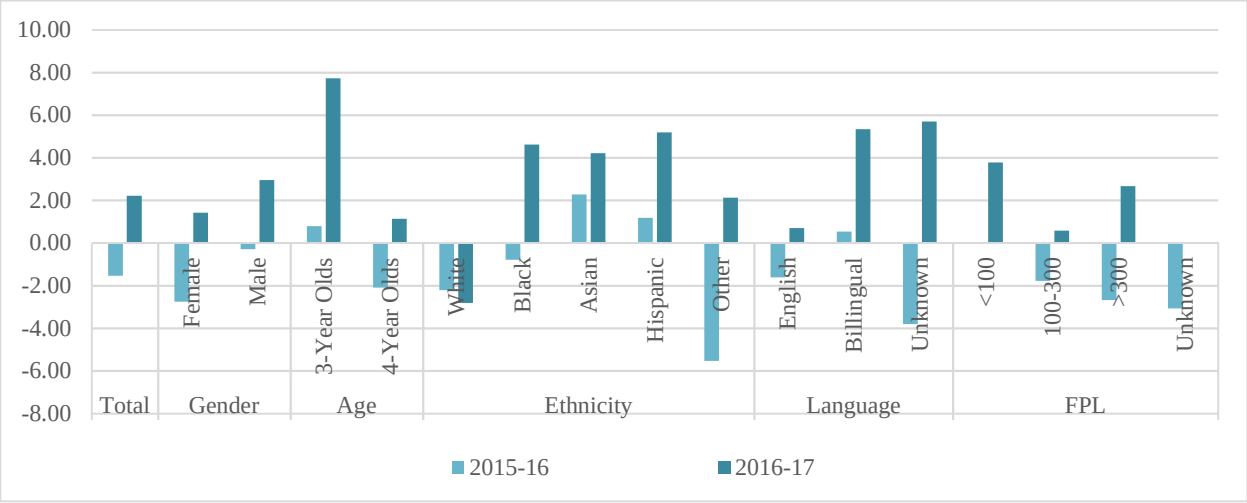


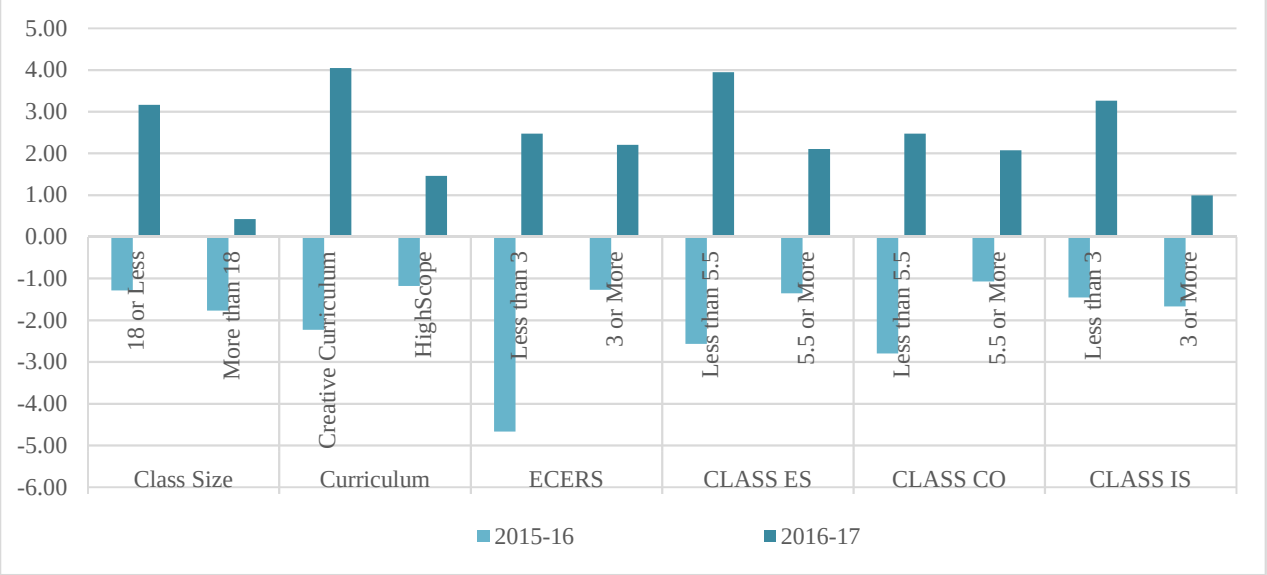
Table 21 reports children's pre- and post-test standardized math scores and gains by selected center characteristics (raw scores are reported in Appendix B.1, Table B.1.6). Again, there is some variation between agencies, and for different quality levels, children in smaller classrooms gain significantly more in this measure. This is also the case for children in classrooms with Creative Curriculum. The quality measures are not associated with gains in the expected direction. The quality measures are not associated with gains in the expected direction.

Table 16. Math means and gains by center characteristics

		N	WJ-AP 2016 Fall		WJ-AP 2017 Spring		WJ-AP Gains 2016–17	
			Mean	SD	Mean	SD	Mean	SD
Total		286	102.37	14.90	104.58	12.89	2.21	12.28
Class Size	18 or Less	187	100.34	15.61	103.51	13.16	3.16	12.75
	More than 18	99	106.19	12.66	106.62	12.18	0.42	11.18
Curriculum	Creative Curr.	84	99.50	17.16	103.54	14.70	4.04	14.49
	HighScope	202	103.56	13.73	105.02	12.07	1.46	11.19
ECERS	Less than 3	15	95.93	14.67	98.47	11.85	2.47	7.79
	3 or More	271	102.72	14.86	104.92	12.88	2.20	12.49
CLASS ES	Less than 5.5	17	95.41	14.64	99.41	13.04	3.94	7.96
	5.5 or More	269	102.81	14.83	104.91	12.84	2.10	12.51
CLASS CO	Less than 5.5	103	99.97	14.46	102.45	12.97	2.47	12.94
	5.5 or More	183	103.72	15.01	105.79	12.72	2.07	11.92
CLASS IS	Less than 3	154	101.08	14.77	104.34	13.75	3.26	13.17
	3 or More	132	103.87	14.97	104.86	11.86	0.99	11.08

Figure 22 shows average scores by center characteristics across years. Again, the most remarkable difference is that all types of centers had increases in standard score gains this year, which is a strong reversal on the trends from the previous year.

Figure 17. Math gains by center characteristics and year



Executive functions

We used two measures of executive functions. The DCCS is an attention shifting test which taps into a child's short-term memory. Table 23 shows children's pre- and post-test DCCS scores by selected child characteristics. The SPP sample gained 0.21 on the DCCS which is equivalent to 0.32 standard deviations, a meaningful change. All subgroups of children evidenced gains between fall and spring. No norms exist for the DCCS. As a reference, the Learning-Related

Cognitive Self-Regulation School Readiness Measures for Preschool Children Study (aka the Self-Regulation Measurement Study) (Meador, et. al, 2013) tested alternative measures of executive functions and included the DCCS. The authors found average DCCS scores of 1.42 at 51–53 months and 1.62 at 57–59 months (an average difference of 0.20 between these two ages); ranges which include the average ages at fall and spring testing in this study (53.2 months in the fall and 59.3 in the spring). Children in SPP show similar gain patterns in relation to age with an average gain of 0.21. Gains were slightly larger for females, four-year-olds, and Black children.

Table 23. DCCS means and gains by child characteristics

		N	DCCS 2016 Fall		DCCS 2017 Spring		DCCS Gains 2016–17	
			Mean	SD	Mean	SD	Mean	SD
Total		286	1.50	0.65	1.72	0.67	0.21	0.60
Gender	Female	138	1.48	0.63	1.75	0.62	0.27	0.57
	Male	148	1.53	0.66	1.69	0.71	0.16	0.62
Age	3-Year-Old Cohort	47	1.04	0.59	1.17	0.56	0.13	0.61
	4-Year-Old Cohort	239	1.59	0.62	1.82	0.63	0.23	0.60
Ethnicity	White	61	1.79	0.49	2.05	0.62	0.26	0.60
	Black	65	1.22	0.67	1.51	0.71	0.29	0.70
	Asian	48	1.52	0.58	1.65	0.56	0.13	0.53
	Hispanic	21	1.33	0.66	1.52	0.60	0.19	0.51
	Other	91	1.55	0.67	1.73	0.65	0.18	0.57
Language	English	196	1.59	0.65	1.82	0.66	0.22	0.60
	DLLs	46	1.41	0.69	1.61	0.61	0.20	0.62
	Unknown	44	1.20	0.51	1.39	0.62	0.18	0.58
FPL	<100	59	1.25	0.68	1.49	0.68	0.24	0.65
	100-300	93	1.74	0.62	1.94	0.69	0.19	0.61
	>300	134	1.45	0.60	1.66	0.60	0.22	0.57

Gains in the DCCS by child characteristics in relation to the previous year are illustrated in Figure 9. Overall gains were slightly higher (0.21 this year versus 0.17 the previous year). There are differences across groups. Gains appear to be driven by gains by females, White, Black children, and children under 100% FPL.

Figure 18. DCCS gains by child characteristics and year

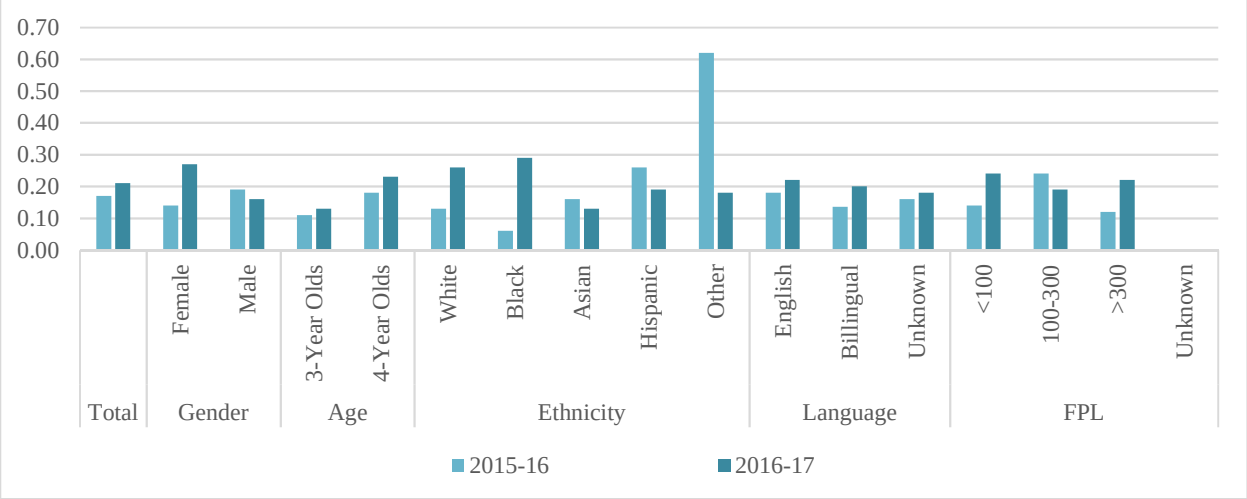


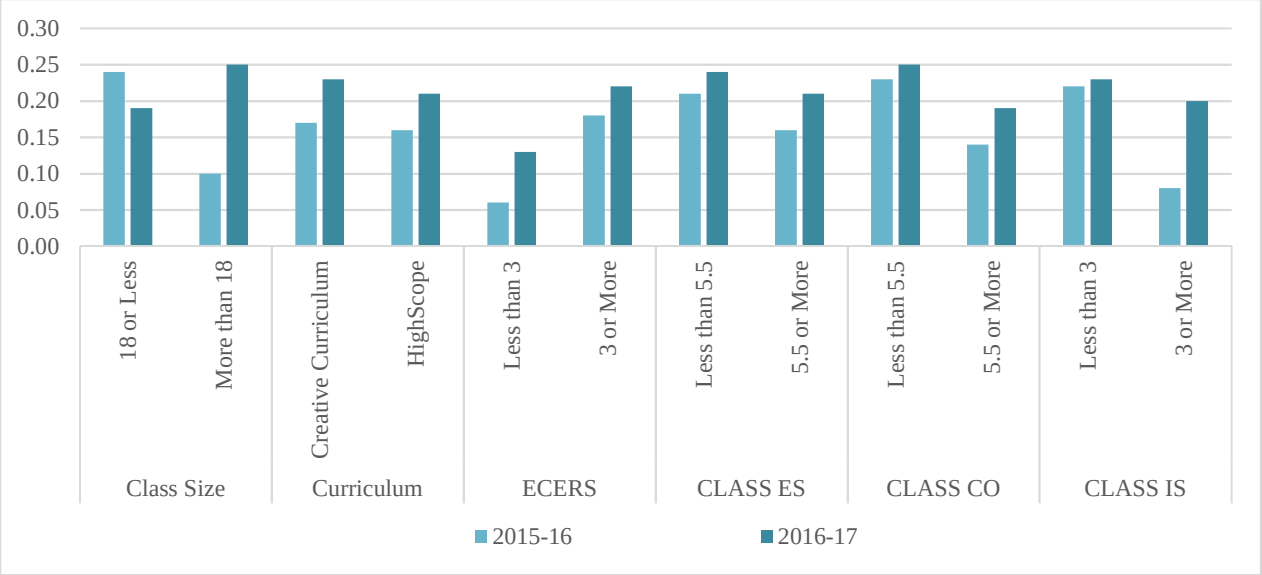
Table 24 presents children's pre- and post-test DCCS scores by selected center characteristics. There are apparent differences in gain between agencies. Differences in gains by curriculum are very minimal. Gains on the DCCS do not differ in the same way across the two quality measures.

Table 24. DCCS means and gains by center characteristics

		N	DCCS 2016 Fall		DCCS 2017 Spring		DCCS Gains 2016–17	
			Mean	SD	Mean	SD	Mean	SD
Total		286	1.50	0.65	1.72	0.67	0.21	0.60
Class Size	18 or Less	187	1.41	0.67	1.60	0.65	0.19	0.60
	More than 18	99	1.69	0.57	1.94	0.64	0.25	0.59
Curriculum	Creative Curr.	84	1.31	0.71	1.54	0.67	0.23	0.59
	HighScope	202	1.58	0.60	1.79	0.65	0.21	0.60
ECERS	Less than 3	15	1.60	0.63	1.73	0.59	0.13	0.64
	3 or More	271	1.50	0.65	1.72	0.67	0.22	0.60
CLASS ES	Less than 5.5	17	1.53	0.62	1.76	0.66	0.24	0.66
	5.5 or More	269	1.50	0.65	1.71	0.67	0.21	0.60
CLASS CO	Less than 5.5	103	1.43	0.65	1.68	0.76	0.25	0.68
	5.5 or More	183	1.55	0.64	1.74	0.61	0.19	0.55
CLASS IS	Less than 3	154	1.45	0.67	1.68	0.65	0.23	0.63
	3 or More	132	1.56	0.62	1.76	0.68	0.20	0.56

Figure 19 illustrates gains in the DCCS in relation to the previous year, by center characteristics. All types of centers in terms of the curriculum chose, and the classroom quality observed, evidenced either similar or larger gains in the DCCS this year.

Figure 19. DCCS gains by center characteristics



In addition to the DCCS, children were assessed with the Peg Tapping measure. Peg Tapping is a measure of inhibitory control. Table 25 shows children's pre- and post-test Peg

Tapping scores by selected child characteristics. No norms exist for this measure, either. Children in SPP across all subgroups gained between fall and spring of the school year, with an overall gain of 2.33 for the full sample (0.40 standard deviations). The Self-Regulation Measurement Study (Meador, et. al, 2013) also included this measure and authors reported average scores of 6.02 at 51–53 months and 8.80 at 57–59 months, with a difference of 2.78. SPP children advanced similarly throughout the preschool year. Among the different subgroups, males, Asian, and children from higher income families gained more than their peers.

Table 25. Peg Tapping means and gains by child characteristics

		N	PT 2016 Fall		PT 2017 Spring		PT Gains 2016–17	
			Mean	SD	Mean	SD	Mean	SD
Total		285	6.33	5.80	8.66	6.05	2.33	5.35
Gender	Female	137	6.29	5.84	8.44	5.98	2.15	5.66
	Male	148	6.37	5.77	8.87	6.13	2.50	5.06
Age	3-Year-Old Cohort	46	0.78	3.15	2.74	4.57	1.96	5.06
	4-Year-Old Cohort	239	7.40	5.58	9.80	5.63	2.40	5.41
Ethnicity	White	61	8.80	5.15	10.97	5.31	2.16	5.51
	Black	64	3.75	5.07	6.11	5.68	2.36	5.29
	Asian	48	6.58	6.52	9.42	6.20	2.83	6.13
	Hispanic	21	5.24	6.02	6.81	6.65	1.57	4.11
	Other	91	6.62	5.55	8.95	5.91	2.33	5.18
Language	English	195	7.19	5.63	9.52	5.67	2.33	5.31
	DLLs	46	5.04	5.77	7.35	6.37	2.30	6.21
	Unknown	44	3.89	5.72	6.25	6.59	2.36	4.64
FPL	<100	58	4.67	5.65	5.29	5.89	0.62	4.14
	100-300	93	8.34	5.67	11.15	5.34	2.81	5.54
	>300	134	5.66	5.59	8.40	5.86	2.74	5.56

Figure 20 illustrates gains in PT by year. Overall, and across all groups except Blacks, gains were slightly larger in 2015–16 than in the current sample of children.

Figure 20. Peg Tapping gains by child characteristics and year

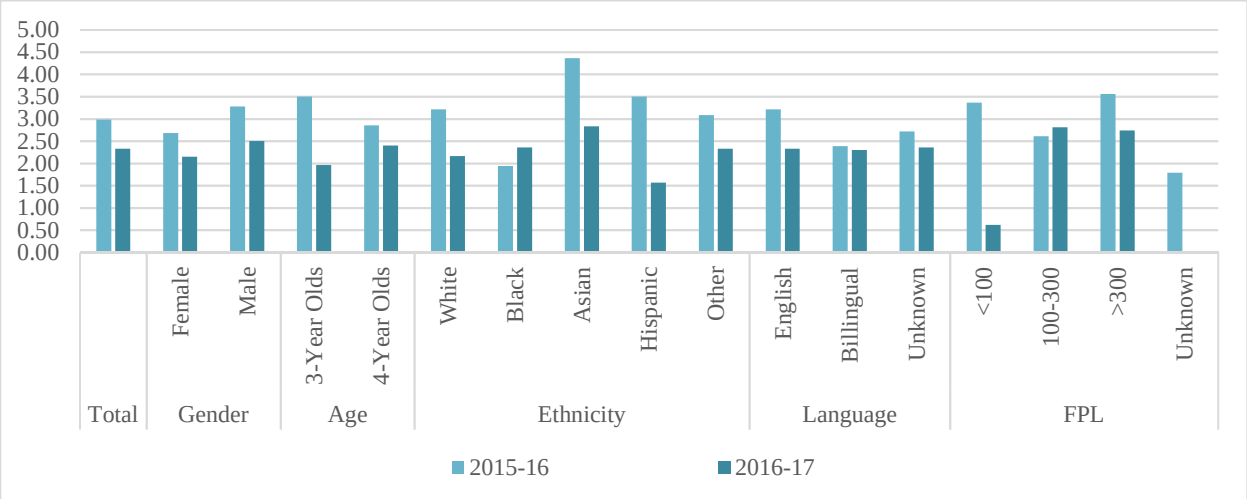


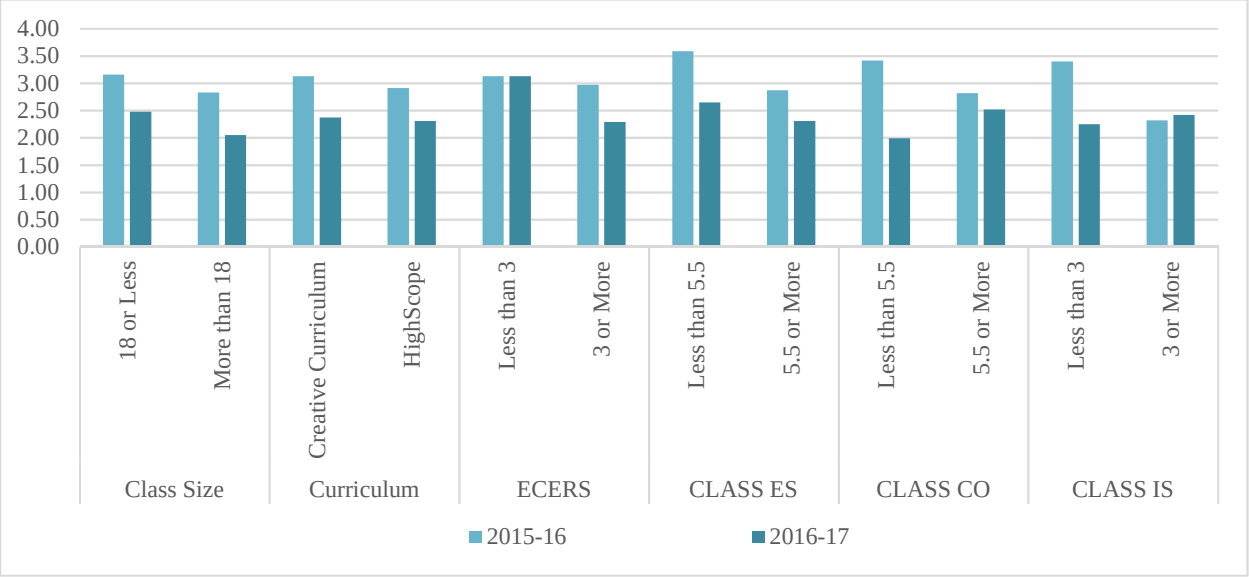
Table 26 shows pre- and post-test Peg-Tapping scores for children in the sample across selected center characteristics. Again, there is variation across agencies, with mean gains varying between 1.72 and 4.35 across these. Higher levels of CLASS CO and IS are associated with higher gains.

Table 26. Peg-Tapping means and gains by center characteristics

		N	PT 2016 Fall		PT 2017 Spring		PT Gains 2016–17	
			Mean	SD	Mean	SD	Mean	SD
Total		285	6.33	5.80	8.66	6.05	2.33	5.35
Class Size	18 or Less	186	5.56	5.72	8.04	6.31	2.48	5.18
	More than 18	99	7.78	5.69	9.83	5.37	2.05	5.66
Curriculum	Creative Curr.	83	5.10	6.06	7.47	6.47	2.37	4.68
	HighScope	202	6.84	5.62	9.15	5.82	2.31	5.61
ECERS	Less than 3	15	7.07	6.39	10.20	5.89	3.13	4.98
	3 or More	270	6.29	5.77	8.58	6.06	2.29	5.37
CLASS ES	Less than 5.5	17	6.00	7.06	8.65	6.06	2.65	6.76
	5.5 or More	268	6.35	5.72	8.66	6.06	2.31	5.26
CLASS CO	Less than 5.5	103	6.10	6.23	8.09	6.23	1.99	4.99
	5.5 or More	182	6.47	5.55	8.99	5.94	2.52	5.55
CLASS IS	Less than 3	154	6.13	6.16	8.38	6.13	2.25	5.25
	3 or More	131	6.57	5.35	8.99	5.96	2.42	5.49

PT gains by center characteristics and year are shown in Figure 21. Gains were slightly smaller for all types of centers across the different characteristics.

Figure 21. Peg-Tapping gains by center characteristics and year



Demographics collected by DEEL differed from those collected through the family survey for 19% (N=88 children). Of these, 32 were children identified as Hispanic/Latinos by DEEL but who were reported as Multi-Racial in the family survey. Consequently, we also calculated gains by race and language groups based on the DEEL indicator (Appendix B.2).

Results for children identified as Hispanic to DEEL differed somewhat from those children identified as Hispanic by the family survey. In particular, the DEEL Hispanic group had stronger gains in receptive vocabulary and executive functions, but weaker gains in literacy and math.

Returning Children

A small subgroup of children in the SPP sample were returning from previous year (45 children, 35 for which we have pre- and post-tests). With the caveat that the N is small, growth for these children is compared to growth for the overall SPP sample in Table 27 below. Returners started the AY with higher scores in most areas relative to the overall sample, and made slightly stronger gains in standard scores than the average SPP child in math and in both measures of executive functions. The lower gains in vocabulary and literacy could be explained by either differentiation not occurring in the classroom (teachers teaching to the bottom) in these two areas while some differentiation occurring in math, and with executive functions strengthening in this second year. Tests of statistical significance comparing returners and non-returners found no differences for all pre- and post-test distributions. Gains in scores were statistically significantly different for PPVT and WJ-LW only, while not so for the rest of the outcomes (in bold, P-Value<0.05).

Table 27. Returners means and gains in standard scores in comparison to all children

	N	2016 Fall		2017 Spring		Gains 2016–17	
		Mean	SD	Mean	SD	Mean	SD
PPVT Non-returners	253	100.60	17.91	103.68	16.95	3.08	9.26
PPVT Returners	35	101.26	18.74	99.77	15.85	**-1.49**	10.92
WJ-LW Non-returners	251	100.23	15.67	101.80	15.16	1.57	9.97
WJ-LW Returners	35	105.49	17.00	102.97	17.22	**-2.51**	10.13
WJ-AP Non-returners	251	102.36	14.80	104.41	12.81	2.04	12.44
WJ-AP Returners	35	102.40	15.85	105.86	13.60	3.46	11.15
DCCS Non-returners	251	1.52	0.65	1.72	0.67	0.19	0.61
DCCS Returners	35	1.40	0.60	1.71	0.67	0.31	0.53
PT Non-returners	250	6.38	5.80	8.64	6.00	2.25	5.36
PT Returners	35	5.97	5.82	8.86	6.51	2.89	5.31

SPP sample in the study versus the rest of SPP children

This study randomly selected children from classrooms, prioritizing children that entered the program through the DEEL enrollment. However, as a complement to this study, Cultivate Learning collected in a separate work with DEEL PPVT on the rest of the children enrolled in the SPP program that did not make part of the sample. This provides an opportunity to compare children in the SPP study sample to the rest of the SPP children in terms of their demographics, as well as their PPVT fall, spring and gain scores (Table 28). Children were comparable, with no statistical difference in gender, age, the percentage of White, Black Hispanic, Asian, and PPVT scores and gains. They however did differ in language for the prevalence of English, Spanish, Chinese non-Mandarin non-Cantonese, and Somali in the sample (no Somali speaking children were in the non-study sample). They also differed in the percentage of 'Other' for race/ethnicity which included the children identified as multi-racial. Differences in proportions that are statistically significant (P<0.05) are in bold. No differences were detected in pre-tests, post-tests nor gains, whether standard nor raw.

Table 28. Study versus non-study SPP children by demographics and PPVT

	SPP study sample		Non-Study Children	
	N	%	N	%
Gender				
Female	141	48.45%	92	45.54%
Male	150	51.55%	110	54.46%
	291	100.00%	202	100.00%
Age at Pre-Test				
3-year-olds	47	16.15%	29	14.36%
4-year-olds	244	83.85%	173	85.64%
	291	100.00%	202	100.00%
Race/Ethnicity				
White	62	21.31%	45	22.28%
Black	67	23.02%	52	25.74%
Asian	48	16.49%	41	20.30%
Hispanic	23	7.90%	23	11.39%
Other	91	**31.27%**	41	**20.30%**
	291	100.00%	202	100.00%
Primary Language				
English	199	**68.38%**	118	**58.42%**
Spanish	3	**1.03%**	8	**3.96%**
Vietnamese	12	4.12%	7	3.47%
Amharic	7	2.41%	4	1.98%
Chinese-Mandarin	4	1.37%	6	2.97%
Chinese-Cantonese	5	1.72%	8	3.96%
Chinese-Other	0	**0.00%**	4	**1.98%**
Somali	4	**1.37%**	0	**0.00%**
Oromo	2	0.69%	3	1.49%
Other	9	3.09%	6	2.97%
Unknown	46	15.81%	38	18.81%
	291	100.00%	202	100.00%
PPVT Standard Scores	Mean (SD) N=288		Mean (SD) N=197	
Fall 2016	100.68	(17.98)	98.73	(19.05)
Spring 2017	103.21	(16.85)	102.02	(17.59)
Gains	2.53	(9.57)	3.03	(10.1)
PPVT Raw Scores				
Fall 2016	71.96	(1.60)	71.96	(1.92)
Spring 2017	85.13	(1.53)	83.96	(1.86)
Gains	12.18	(.75)	12.87	(.86)

Multivariate Analyses

Through multivariate analyses we examine the association between children's learning gains and program features while simultaneously controlling for children's characteristics. In addition, we

are also able to examine the contribution of the program in relation to experiences of other children in the City of Seattle in a separate set of estimations. We include information on the age of children, gender, race and ethnicity, home language, income, and FPL. Program features for SPP children include class size, agency and classroom quality. The analyses also take into account that scores of children who are in classrooms together cannot be considered to be independent of each other.

The first set of models assess the association between SPP children's learning gains, their characteristics and program features. We conduct separate analyses with the two measures of quality, one controlling for quality as measured by the ECERS-3, and the other for quality as measured by the CLASS.

Table 29 & 30 present the estimates of the associations of program features and child characteristics with children's development. Table 29 includes these with ECERS-3 as the measure of classroom quality and Table 30 does so with CLASS domains. In these estimations, we only examine the association between the child characteristics and the center characteristics to the development of children. Statistically significant results are highlighted in bold.

As for classroom features, no association exists between the classroom size and children's performance, accounting for all individual and other program features measured. ECERS-3 and children's performance is only significantly associated for literacy raw score (see Appendix Tables C.1.1 and C.1.2 for raw score estimations). Unexpectedly, CLASS emotional supports score is negatively associated with DCCS gains. Blacks and Hispanics evidenced lower receptive vocabulary, while children categorized as Other evidenced lower literacy scores (see Table 30). No systematic differences were evidenced by income. Children's outcomes did differ for girls (were lower), but only for math, in the model with CLASS dimensions. Children under 100% FPL scored lower in the Peg Tapping task than children above 300% FPL (the omitted group). Agency-selected children had higher gains in receptive vocabulary alone. There are no consistent patterns of advantages or disadvantages due to children's characteristics across all the areas of development measured on children that emerge from these results.

Table 29. Multivariate analyses of children's 2016–17 standard score gains in relation to child and site or classroom characteristics and ECERS-3

Variables	Rec. Vocabulary (PPVT/TVIP)	Literacy (WJ/WM-LW)	Math (WJ/WM-AP)	Executive Function	
				DCCS	PT
Female	1.303	0.330	-2.248	0.083	-0.571
	(0.98)	(1.05)	(1.16)	(0.06)	(0.54)
Black	**-4.572***	-0.606	-1.466	-0.028	-0.378
	(1.90)	(1.95)	(2.23)	(0.11)	(1.00)
Asian	-1.374	0.817	-0.603	-0.149	1.263
	(1.95)	(2.00)	(2.23)	(0.11)	(1.02)
Hispanic	**-5.642***	-4.454	-2.217	-0.189	-1.511
	(2.37)	(2.52)	(2.81)	(0.14)	(1.25)
Other Race	-0.460	**-3.166***	0.769	-0.153	-0.275
	(1.49)	(1.56)	(1.75)	(0.09)	(0.79)
DLL	0.102	0.102	0.569	0.007	-0.732
	(1.68)	(1.77)	(1.96)	(0.10)	(0.90)
Agency Selected	**5.206***	1.150	1.823	-0.066	1.980
	(2.51)	(2.68)	(2.95)	(0.15)	(1.38)
Income ≤20k	-2.741	-1.670	-1.389	0.032	-1.481
	(2.37)	(2.52)	(2.82)	(0.14)	(1.29)
Income 21K-40K	-1.290	-0.469	-0.928	0.004	-1.457
	(2.09)	(2.23)	(2.47)	(0.13)	(1.14)
Income 41K-60K	1.859	-1.962	-1.143	0.085	-0.477
	(2.13)	(2.28)	(2.54)	(0.13)	(1.17)
Income 61K-80K	**-3.660***	1.869	-1.692	0.039	0.175
	(1.77)	(1.88)	(2.08)	(0.11)	(0.96)
FPL <100%	0.390	3.951	-4.031	-0.036	**-2.520***
	(2.15)	(2.27)	(2.51)	(0.13)	(1.16)
FPL 100-300%	0.009	1.772	-1.464	-0.055	-0.212
	(1.68)	(1.79)	(1.98)	(0.10)	(0.91)
Class Size	0.335	-0.133	0.417	0.008	0.179
	(0.27)	(0.29)	(0.32)	(0.02)	(0.15)
ECERS	0.345	2.210	2.857	-0.114	-0.678
	(1.44)	(1.53)	(1.68)	(0.09)	(0.77)
N	288	286	286	286	285

* p<0.05; ** p<0.01; *** p<0.001. Note: Reference groups omitted from the estimation are Males, White, English, FPL 300%+ and Income>80 thousand. Other controls are pre-test, age in months, days between tests, agencies, and an indicator for missing language or income. Standardized scores are used for PPVT, and WJ or WM. Errors are clustered by site.

Table 30. Multivariate analyses of children's 2016–17 standard score gains in relation to child and site or classroom characteristics CLASS dimensions

Variables	Rec. Vocabulary (PPVT/TVIP)	Literacy (WJ/WM-LW)	Math (WJ/WM-AP)	Executive Function	
				DCCS	PT
Females	1.121	0.225	**-2.373***	0.080	-0.610
	(0.97)	(1.05)	(1.17)	(0.06)	(0.53)
Black	**-4.161***	-0.803	-1.434	-0.022	-0.316
	(1.88)	(1.96)	(2.25)	(0.11)	(1.00)
Asian	-1.561	0.650	-0.698	-0.154	1.205
	(1.94)	(1.99)	(2.24)	(0.11)	(1.01)
Hispanic	**-5.422***	-4.444	-2.224	-0.174	-1.410
	(2.34)	(2.52)	(2.83)	(0.14)	(1.25)
Other Race	-0.449	**-3.395***	0.765	-0.162	-0.326
	(1.48)	(1.56)	(1.76)	(0.09)	(0.79)
DLL	-0.037	0.559	0.802	0.013	-0.676
	(1.66)	(1.76)	(1.97)	(0.10)	(0.90)
Agency Selected	**5.193***	1.086	1.443	-0.035	2.137
	(2.49)	(2.68)	(2.98)	(0.15)	(1.37)
Income ≤20k	-3.158	-1.347	-1.376	0.044	-1.395
	(2.35)	(2.52)	(2.85)	(0.14)	(1.29)
Income 21K-40K	-1.486	-0.283	-1.004	0.018	-1.364
	(2.07)	(2.23)	(2.49)	(0.13)	(1.14)
Income 41K-60K	1.773	-1.776	-1.051	0.093	-0.404
	(2.12)	(2.29)	(2.57)	(0.13)	(1.16)
Income 61K-80K	**-3.727***	2.076	-1.683	0.049	0.244
	(1.75)	(1.88)	(2.09)	(0.11)	(0.96)
FPL <100%	1.184	3.884	-3.688	-0.048	**-2.577***
	(2.13)	(2.27)	(2.53)	(0.13)	(1.16)
FPL 100-300%	0.340	1.773	-1.158	-0.076	-0.344
	(1.66)	(1.79)	(2.00)	(0.10)	(0.91)
Class Size	0.293	-0.269	0.296	0.010	0.198
	(0.29)	(0.31)	(0.35)	(0.02)	(0.16)
CLASS ES	-1.914	-0.825	0.114	**-0.155***	-1.166
	(1.25)	(1.33)	(1.48)	(0.07)	(0.67)
CLASS CO	1.853	0.780	0.819	0.041	0.276
	(1.19)	(1.31)	(1.45)	(0.07)	(0.66)
CLASS IS	0.956	-1.542	-0.010	-0.065	-0.364
	(0.81)	(0.88)	(0.98)	(0.05)	(0.45)
N	288	286	286	286	285

* p<0.05; ** p<0.01; *** p<0.001. Note: Reference groups omitted from the estimation are Males, White, English, FPL 300%+ and Income>80 thousand. Other controls are pre-test, age in months, days between tests, agencies, and an indicator for missing language or income. Standardized scores are used for PPVT, and WJ or WM. Errors are clustered by site.

A second set of models compare the SPP sample (Group A) to the recruited sample of other children in the City of Seattle. The 153 children recruited are grouped into two groups: those recruited from the waiting list (Group B) and those recruited from centers in which children on the waiting list were attending (Group C). Table 31 illustrates demographic differences and/or similarities across the SPP sample and these two groups, only for children with pre- and post-tests. Children enrolled from the waiting list (Group B) were more likely to be females, more likely to be 3-year-olds, more likely to be dual language learners, similar in

income reported (for those with income known), and more likely to be White or Asian. Children enrolled from centers (Group C) were similarly balanced by gender, and more closely balanced by age, but more likely to be English speaking as well, with incomes above 80,000 per year and White.

Analyses to assess baseline equivalence on observable variables (including pre-test, race, income, and gender) found no statistical differences[7] in gender and income (when distinguishing above versus below $80,000)[8] between group A and B, and statistical differences in age cohorts, home language, race/ethnicity and income when distinguishing among levels under $80,000. In addition, there are no statistical differences in gender and language between group A and groups B and C pooled, but there are statistical differences in age cohorts, race/ethnicity and income. Quality of classrooms for Group C (N=7) is shown in Section 5, above (Tables 14 and 15). Classroom quality experienced by this group is lower on average for both the ECERS-3 and all three dimensions of the CLASS despite this group being more likely to earn above $81,000 a year and more likely to be White. Statistically significant differences in proportions relative to the SPP group (P<0.05) are in bold.

Table 31. Child demographics for SPP children relative to children in Groups B and C

Child Characteristics	SPP Children 2016–17		Waiting List Comparison (Group B)		Center Based Comparison (Group C)		Pooled Comparisons (Group B+C)	
	N	%	N	%	N	%	N	%
Gender								
Female	141	48.45%	36	59.02%	36	45.00%	72	51.06%
Male	150	51.55%	25	40.98%	44	55.00%	69	48.94%
Age at Pre-Test								
3-Year-Olds	47	16.15%	27	**44.26%**	24	**30.00%**	51	**36.17%**
4-Year-Olds	244	83.85%	34	**55.74%**	56	**70.00%**	90	**63.83%**
Primary Language								
English	199	68.38%	44	**72.13%**	65	**81.25%**	109	77.30%
Dual Language	46	15.8%	16	**26.6%**	1	**1.25%**	17	12.06%
Unknown	46	15.81%	1	**1.64%**	14	**17.50%**	15	10.64%
Income								
20,000 or Less	43	14.78%	7	**11.48%**	-	-	7	**4.96%**
21,000-40,000	54	18.56%	18	**29.51%**	4	**5.00%**	22	**15.60%**
41,000-60,000	42	14.43%	5	**8.20%**	3	**3.75%**	8	**5.67%**
61,000-80,000	44	15.12%	12	**19.67%**	6	**7.50%**	18	**12.77%**
81,000 or more	59	20.27%	16	26.23%	52	**65.00%**	68	**48.23%**
Unknown	49	16.84%	3	4.92%	15	**18.75%**	18	**12.77%**
FPL Percentage								
Less than 100%	61	20.96%	2	3.28%	-	-	2	1.42%
100 – 199%	68	23.37%	-	-	-	-	-	-
200 – 299%	67	23.02%	-	-	-	-	-	-
>300%	95	32.65%	-	-	-	-	-	-
Unknown	-	-	59	96.72%	80	100%	139	98.58%

[7] Measured at a 5% difference.
[8] For a family of four, $72,750 was the 2016 300% FPL.

Race/Ethnicity								
White	62	21.31%	25	**40.98%**	52	**65.00%**	77	**54.61%**
Black	67	23.02%	5	**8.20%**	-	-	5	**3.55%**
Asian	48	16.49%	15	**24.59%**	4	**5.00%**	19	**13.48%**
Hispanic	23	7.90%	4	**6.56%**	3	**3.75%**	7	**4.96%**
Multi-Racial	83	28.52%	11	**18.03%**	7	**8.75%**	18	**12.77%**
Other	8	2.75%	-	-	-	-	-	-
Unknown	-	-	1	**1.64%**	14	**17.50%**	15	**10.64%**

In Table 32 we examine this contribution in relation to the experiences of other children in the City of Seattle. For these main estimates, the comparison is to the group recruited from the waiting list, Group B. Main analyses do not include program features as these are not available for most children in the comparison groups. Results show positive but non-significant differences in gains relative to waiting list children in vocabulary, literacy and math standard scores, and non-significant negative gains in the DCCS. Significant negative effects are observed for the PT measure of executive functions.

Table 32. Multivariate analyses of SPP children's 2016–17 gains in relation to children in the waiting list

Variables	**Rec. Vocabulary (PPVT/TVIP)**	**Literacy (WJ/WM-LW)**	**Math (WJ/WM-AP)**	**Executive Function**	
				DCCS	**PT**
SPP program	1.561	2.105	1.046	-0.130	**-1.314****
	(1.08)	(1.09)	(2.28)	(0.09)	(0.45)
Female	0.737	0.252	-1.995	0.095	-0.098
	(0.72)	(0.90)	(1.23)	(0.06)	(0.62)
Black	**-6.268*****	0.043	-2.832	-0.108	-0.009
	(1.55)	(1.87)	(2.35)	(0.09)	(0.81)
Asian	-3.502	-0.291	-1.247	-0.125	1.413
	(2.22)	(2.25)	(2.72)	(0.08)	(0.86)
Hispanic	**-5.974****	-3.757	-2.538	-0.217	-0.390
	(1.89)	(2.03)	(2.39)	(0.11)	(0.86)
Other Race	-2.340	-1.957	1.189	-0.177*	0.056
	(1.27)	(1.40)	(1.46)	(0.08)	(0.65)
DLL	-0.689	2.751	0.664	-0.059	-0.855
	(1.90)	(2.28)	(2.13)	(0.07)	(0.76)
Agency Selected	1.170	0.122	0.785	-0.129	0.630
	(1.45)	(1.42)	(1.10)	(0.10)	(0.61)
Income ≤20k	**-3.954***	-1.204	-3.339	-0.060	**-3.406*****
	(1.90)	(2.08)	(2.24)	(0.14)	(0.57)
Income 21K-40K	-1.760	-1.173	-1.924	0.045	**-1.641***
	(1.73)	(1.89)	(1.91)	(0.11)	(0.68)
Income 41K-60K	1.406	-1.534	-2.257	-0.075	-1.404
	(1.63)	(1.53)	(2.21)	(0.13)	(0.88)
Income 61K-80K	**-5.466***	0.797	-2.744	-0.042	-0.698
	(2.27)	(1.74)	(1.37)	(0.10)	(0.84)
N	347	346	346	345	344

* p<0.05; ** p<0.01; *** p<0.001. Note: Reference groups omitted from the estimation are Males, White, English, and Income>80 thousand. FPL information not available for children not in SPP, so this variable was excluded. Other controls are pre-test, age in months, days between tests, and an indicator for missing language, race (cases in control group only) or income. Standardized scores are used for PPVT, and WJ or WM. Errors are clustered by site.

Sensitivity Analyses

Two types of sensitivity analyses were conducted to assess the robustness of our findings for the main analyses of features of classrooms, or characteristics of children, that were related to their development. First, we repeated the analyses with raw scores because imperfections in the standardization could affect results. Second, we investigated whether a quality threshold made a difference.

The results of the three types of sensitivity analyses are summarized as follows.

(1) Results of analyses on raw scores for the PPVT, LW and AP measures (Tables C.1 using ECERS and C.2 using CLASS) are consistent with the standard score analyses. The exception is that ECERS-3 does evidence an effect in literacy gains for SPP children.

(2) Analyses investigating thresholds of quality are reported in Appendix Tables C.3 for ECERS and C.4 for CLASS.[9] We find that no association between the ECERS-3 threshold above 3 and children's standard score gains (or raw score gains, either, although these are not reported). We observe a positive association for CLASS CO levels above 5.5 with literacy scores.

In addition, we ran five additional analyses on the estimations of SPP children's development in relation to other children in Seattle (Table 33). The first model includes the children recruited through the waiting list (Group B), but also pools the children recruited in centers attended by waiting list children (Group C). Model 2 replicates the main analyses in Table 32 with raw scores. Model 3 replicates model 1 with raw scores. Model 4 pools the waiting list children (Groups B) with only a subgroup of children that had a higher probability of being a SPP attender given their socio-demographic characteristics.[10] Model 5 replicates Model 4 with raw scores. All estimations show differences between the SPP groups and the different comparison groups as expressed in effect sizes (as a fraction of a standard deviation of the norm in estimations with standard scores, or as a fraction of the standard deviation of the SPP children in their fall scores in estimations with raw scores). The differences between estimations are outlined under each model. Overall, we find consistently positive differences relative to the comparison groups of children compared in receptive vocabulary and literacy, effects varying for math, depending on whether looking at standard or raw scores, and negative effects in executive functions. Across all these, only the negative difference in Peg Tapping was statistically significant and in Model 3, Literacy was also statistically significant, which speaks to this effect being positive but on the margin of significance across models. Even though the comparisons groups were more likely to be White, and higher income as reported, children in SPP classrooms gained quite similarly across most areas to children in the comparison groups.

[9] Burchinal et al. (2010) found evidence of CLASS IS thresholds at 3.25, and CLASS ES in the 5-7 range, and Hatfield et al. (2016) found evidence of CLASS IS threshold at 3 and CLASS ES and CO at 6. Given the distributions of quality in the sample, we chose to use a level of 3 for the ECERS and levels of 5.5 for CLASS emotional support and classroom organization scales, and a level of 3 for CLASS instructional supports.

[10] We estimated the probability of SPP participation for all pooled children (Groups A, B, C) in relation to their socio-demographic characteristics (age, gender, race and ethnicity, language and income) and the calculated their probability scores. We then limited this sample to only children with a probability of being an SPP attender $P \geq 0.5$.

Table 33. Multivariate analyses of children's 2016–17 gains in relation to child and the comparison group

Variables	Model Table 20	Model 1	Model 2	Model 3	Model 4	Model 5
Rec. Vocabulary (PPVT/TVIP)	0.10	0.05	0.06	0.04	0.09	0.06
Literacy (WJ/WM-LW)	0.14	0.11	0.11	**0.11**	0.11	0.11
Math (WJ/WM-AP)	0.07	-0.02	0.04	0.00	0.00	0.00
DCCS	-0.20	-0.14			-0.19	
PT	**-0.23**	**-0.22**			**-0.23**	
With Waiting-list only	yes	-	yes	-	-	-
With Waiting-list (Group B) & center-control (Group C)*	-	yes	-	yes	-	-
With Waiting-list (Group B) & limited center-control (Group C)*	-	-	-	-	yes	yes
Raw scores for PPVT or WJ	-	-	yes	yes	-	yes
N (varies by outcome)	344-347	424-426	346-347	425-426	363-366	365-366

Note: ES for significant associations between SPP and the outcomes are shown in bold font. Note: Controls include age in months, days between tests, gender, race or ethnicity, bilingual, income and FPL, and indicators for missing language, race (cases in control group only) or income. Errors are clustered by site.

5. To what extent are children's learning gains moderated by other learning activities, particularly parent activities and prior center-based care and education?

In the family survey we included questions addressing whether children had attended a center in the previous year, whether parents felt a connection with the preschool (to the teacher, the preschool, receive work samples, receive assessment results, know about the curriculum, feel welcome in the preschool, have received feedback on the child's progress), whether they perceived a positive changes in their children (in language, physical, behavioral/social-emotional, literacy, math, science), whether teachers communicated with them (talks to them each day, uses a curriculum for teaching, teaches behavioral skills, teachers academic skills, tracks child progress, is fluent in child's home language, has a BA, engages in training opportunities), and about their interactions with children (read books, tells stories, sing songs, help do crafts with child, among others). These variables capture aspects related to parental investments in children, as well as a center's interactions and communication with families. In this section we address whether these experiences are related to children's gains in the SPP sample, and further below, we assess the extent to which these moderate the contribution of the program to children's learning and development.

Table 34 below presents similar estimates to those in Table 29 with these additional set of variables. The composite measuring the connection to the preschool reported by the parent is positively associated with children's literacy and DCCS gains. The communication with the

teacher had positive associations with the PT measure of executive functions. Parent's perception of whether their child has shown various positive changes since enrollment in the program was, on the other hand, negatively related to both measures of executive functions; an unusual finding. Having had previous early childhood center-based experiences was positively associated with math and the DCCS. The composite of parent-child interactions measured in this study had no associations with any of the outcomes. ECERS-3 estimates did not vary.

Table 34. SPP children gains including parent perceptions, interactions and previous center-based experience in children with ECERS

Variables	Rec. Vocabulary (PPVT/TVIP)	Literacy (WJ/WM-LW)	Math (WJ/WM-AP)	Executive Function	
				DCCS	PT
Female	1.100	0.622	-2.077	0.090	-0.616
	(0.98)	(1.04)	(1.15)	(0.06)	(0.53)
Black	**-4.765***	-0.565	-1.234	-0.032	-0.406
	(1.90)	(1.94)	(2.20)	(0.11)	(0.98)
Asian	-1.381	0.913	-0.097	-0.114	1.411
	(1.94)	(1.99)	(2.21)	(0.11)	(1.00)
Hispanic	**-5.897***	-4.505	-2.793	-0.171	-1.756
	(2.38)	(2.50)	(2.79)	(0.14)	(1.23)
Other Race	-0.568	**-3.217***	1.071	-0.143	-0.142
	(1.49)	(1.55)	(1.73)	(0.09)	(0.78)
Bilingual	0.188	-0.763	0.754	0.014	-0.471
	(1.74)	(1.81)	(2.00)	(0.10)	(0.92)
Agency Selected	**5.216***	0.627	2.166	-0.039	1.946
	(2.59)	(2.76)	(3.02)	(0.15)	(1.40)
Income <20K	-2.678	-1.831	-0.830	0.025	-1.075
	(2.39)	(2.52)	(2.81)	(0.14)	(1.28)
Income 21K-40K	-1.110	-0.878	-0.930	-0.018	-1.297
	(2.09)	(2.22)	(2.44)	(0.12)	(1.12)
Income 41K-60K	1.973	-2.410	-0.502	0.090	-0.395
	(2.16)	(2.30)	(2.54)	(0.13)	(1.16)
Income 61K-80K	**-3.563***	1.914	-1.182	0.015	0.192
	(1.78)	(1.88)	(2.07)	(0.11)	(0.95)
FPL <100%	0.429	4.304	-3.826	-0.032	**-2.729***
	(2.14)	(2.25)	(2.48)	(0.13)	(1.14)
FPL 100-300%	0.364	1.994	-0.996	-0.016	0.099
	(1.68)	(1.79)	(1.98)	(0.10)	(0.90)
Class Size	0.363	-0.195	0.412	0.007	0.172
	(0.27)	(0.29)	(0.32)	(0.02)	(0.15)
Positive Change	-0.120	0.020	-0.052	**-0.011***	**-0.124***
	(0.09)	(0.10)	(0.11)	(0.01)	(0.05)

Connection to preschool	0.104	0.207	0.090	**0.013***	0.074
	(0.11)	(0.11)	(0.13)	(0.01)	(0.06)
Teacher communication	0.408	-0.711	0.085	0.012	**0.479****
	(0.34)	(0.37)	(0.40)	(0.02)	(0.18)
Interaction with child	0.003	-0.074	-0.077	0.003	-0.029
	(0.05)	(0.05)	(0.05)	(0.00)	(0.03)
Previous center experience	0.447	-0.070	**2.896***	**0.161****	0.324
	(0.97)	(1.04)	(1.14)	(0.06)	(0.52)
ECERS	0.438	1.871	2.455	-0.124	-0.793
	(1.44)	(1.52)	(1.66)	(0.08)	(0.76)
Observations	288	286	286	286	285

* p<0.05; ** p<0.01; *** p<0.001. Note: Controls include age in months, days between tests, gender, race or ethnicity, bilingual, income and FPL, and indicators for missing language or income. Errors are clustered by site.

Table 35 below replicates these estimations with the CLASS. As in Table 34, there are significant negative associations between the parental perceptions composite and the executive functions measure, and positive associations between previous center experience and Math and the DCCS. The negative association between CLASS ES and the DCCS remains.

Table 35. SPP children gains including parent perceptions, interactions and previous center-based experience in children with CLASS

Variables	**Rec. Vocabulary (PPVT/TVIP)**	**Literacy (WJ/WM-LW)**	**Math (WJ/WM-AP)**	**Executive Function**	
				DCCS	**PT**
Female	0.985	0.554	-2.167	0.089	-0.624
	(0.97)	(1.04)	(1.16)	(0.06)	(0.53)
Black	**-4.347***	-0.640	-1.161	-0.022	-0.356
	(1.89)	(1.94)	(2.22)	(0.11)	(0.98)
Asian	-1.484	0.739	-0.126	-0.119	1.400
	(1.93)	(1.98)	(2.22)	(0.11)	(1.00)
Hispanic	**-5.673***	-4.535	-2.809	-0.156	-1.651
	(2.35)	(2.50)	(2.80)	(0.14)	(1.23)
Other Race	-0.497	**-3.351***	1.103	-0.149	-0.195
	(1.48)	(1.55)	(1.74)	(0.09)	(0.78)
Bilingual	0.014	-0.318	0.965	0.015	-0.436
	(1.71)	(1.80)	(2.00)	(0.10)	(0.91)
Agency Selected	**5.185***	0.594	1.917	-0.007	2.223
	(2.58)	(2.77)	(3.06)	(0.15)	(1.41)
Income <20K	-3.102	-1.365	-0.816	0.036	-0.989
	(2.37)	(2.52)	(2.83)	(0.14)	(1.27)
Income 21K-40K	-1.367	-0.649	-1.023	-0.006	-1.209
	(2.07)	(2.21)	(2.45)	(0.12)	(1.12)
Income 41K-60K	1.885	-2.162	-0.444	0.095	-0.396
	(2.14)	(2.30)	(2.56)	(0.13)	(1.16)
Income 61K-80K	**-3.649***	2.146	-1.161	0.022	0.247
	(1.76)	(1.87)	(2.07)	(0.11)	(0.95)
FPL <100%	1.166	4.180	-3.515	-0.045	**-2.845***
	(2.13)	(2.26)	(2.50)	(0.13)	(1.14)
FPL 100-300%	0.684	1.911	-0.678	-0.039	-0.039
	(1.67)	(1.79)	(1.99)	(0.10)	(0.90)
Class Size	0.316	-0.309	0.284	0.009	0.168
	(0.30)	(0.31)	(0.35)	(0.02)	(0.16)

Positive Change	-0.109	0.023	-0.053	**-0.011***	**-0.122***
	(0.09)	(0.10)	(0.11)	(0.01)	(0.05)
Connection to preschool	0.100	0.182	0.084	0.012	0.062
	(0.11)	(0.12)	(0.13)	(0.01)	(0.06)
Teacher communication	0.357	-0.704	0.108	0.008	0.467*
	(0.34)	(0.37)	(0.40)	(0.02)	(0.18)
Interaction with child	0.001	-0.091	-0.085	0.003	-0.032
	(0.05)	(0.05)	(0.06)	(0.00)	(0.03)
Previous center experience	0.574	-0.143	**3.007****	**0.152****	0.253
	(0.96)	(1.04)	(1.15)	(0.06)	(0.52)
CLASS_ES	-1.386	-1.335	0.179	**-0.152***	-0.961
	(1.27)	(1.34)	(1.49)	(0.07)	(0.68)
CLASS_CO	1.640	0.731	0.958	0.032	0.378
	(1.21)	(1.32)	(1.47)	(0.07)	(0.67)
CLASS_IS	1.007	-1.412	0.004	-0.052	-0.499
	(0.82)	(0.89)	(0.99)	(0.05)	(0.45)
Observations	288	286	286	286	285

* p<0.05; ** p<0.01; *** p<0.001. Note: Controls include age in months, days between tests, gender, race or ethnicity, bilingual, income and FPL, and indicators for missing language or income. Errors are clustered by agency.

We also inquired into the association of these variables for children in SPP in relation to children in the control group. Table 36 summarizes these indices and variables for the SPP children as well as the control group included in these estimations. Families in the SPP group did report higher overall positive change, connections with the preschool and teacher communication levels, while reporting lower amounts of interactions with their children and less previous center-based experienced than their counterparts in the control group.

Table 36. Summary statistics for moderating parenting and learning supports

Variable	N	SPP		Control (Group C+)	
		Mean	SD	Mean	SD
Positive change	371	30.87	14.82	24.36	17.09
Connection to preschool	371	29.97	14.09	23.38	16.16
Teacher communication	371	5.14	2.83	3.83	2.96
Interaction with child	371	53.11	25.5	59.59	19.27
Previous center experience	371	0.34	0.78	0.39	0.68

In addition, we replicated Model 4 from Table 33 assessing the association between SPP and children's learning gains incorporating these variables. Table 37 below shows the contribution of the SPP program for children in the program in comparison to other children in the City of Seattle after being moderated for the variables on parent interactions, the connection and communication between parents and centers, parent's perceptions of children's changes and previous center experience.

As in Table 33, the contribution of SPP is positive across receptive vocabulary and literacy. However, after controlling for the different moderators, this is not the case for math anymore. Moreover, negative statistically significant effects are present for both measures of executive functions. These moderators measure various aspects the schools do in terms of reaching and engaging parents in terms of their children's learning, as well as parental investments in children. To the extent that SPP schools do a better job in parent communication

and in connecting with families, these effects are capturing a contribution of the program to children's learning.

Table 37. Multivariate analyses of children's 2016–17 gains in relation to child and the comparison group

Variables	Positive Change	Connection w/Preschool	Teacher Commun.	Interactions w/Child	Previous Center Exp.	All variables
Rec. Vocabulary (PPVT/TVIP)	0.09	0.07	0.07	0.09	0.10	0.07
Literacy (WJ/WM-LW)	0.03	-0.01	0.06	0.11	0.11	-0.02
Math (WJ/WM-AP)	-0.06	-0.07	-0.02	-0.01	0.00	-0.12
DCCS	**-0.24**	**-0.30**	-0.22	-0.18	-0.20	**-0.28**
PT	**-0.20**	**-0.24**	**-0.29**	**-0.25**	**-0.23**	**-0.30**
With Waiting-list (Group B) & limited center-control (Group C)*	yes	yes	yes	yes	yes	yes
All parent perception variables						yes
N (varies by outcome)	363-366	363-366	363-366	363-366	363-366	363-366

* p<0.05; ** p<0.01; *** p<0.001. Note: Controls include age in months, days between tests, gender, race or ethnicity, bilingual, income and FPL, and indicators for missing language or income. Errors are clustered by site.

6. What activities do children engage in, and is there scope for their interests and active participation?

To inquire into whether classrooms offered scope for children's interests and active participation, we focus on specific indicators in the ECERS-3 that expressly address interactions and the ways in which staff actively engage children. We report as bar graphs the frequency with which classrooms met these specific indicators. Indicators are grouped by item in the graphs to which they belong.

In the ECERS-3, indicators are organized in 4 levels: inadequate, minimal, good and excellent. Taking this into account, for each item we scored individual indicators based on the percentage of classrooms that exceeded that indicator. To assist readers in following the indicators we color coded them in the figures that follow. The lowest level indicators--inadequate--are denoted in red. Typically these indicators are phrased negatively so we have reversed scored them. We report the percentage surpassing each level, so when we report 100% for a red indicator that means that none of the classrooms did what is described by the (negative) indicator, and all classrooms scored better than inadequate. Indicators of "minimal" (scores of 3.00) on the ECERS-3 are denoted by yellow bars, and again we report the percentage surpassing that level. For example, 50% means that half of the programs scored at least minimal. Indicators that constitute a score of a 5.00 or "good" were coded blue bars, and for a 7.00 or "excellent" green was used.

This system will help visualize the percentage of classrooms meeting indicators relating to the engagement of children by staff providing initial insights into areas that need improvement

and could be targeted through the continuous quality improvement cycle. Special attention should be given to the blue and green bars, as it is desirable for classrooms to be in the good to excellent range on these which are higher order interactions.[11] The detailed information used to construct these figures is reported in Appendix E.

Figure 22 reports on the percentage of classrooms that met the indicators for Item 5 on staff talking about the display of materials and staff pointing out and reading words, and Item 13, staff encouraging children to use language through responding to them, helping them communicate with each other and talking beyond classroom activities, among other things. The upper bar is the percent of classrooms the met this indicator the previous year, while the lower bar is the percent of classrooms that met it this year. In 2017, there was a stark increase in the percentage of classrooms where staff pointed out to displayed words and read these out loud and on the percentage of staff-child conversations going beyond activities planned in the classroom.

Figure 22. Indicators met on display and use of language (N=32 in 2017, N=14 in 2016)

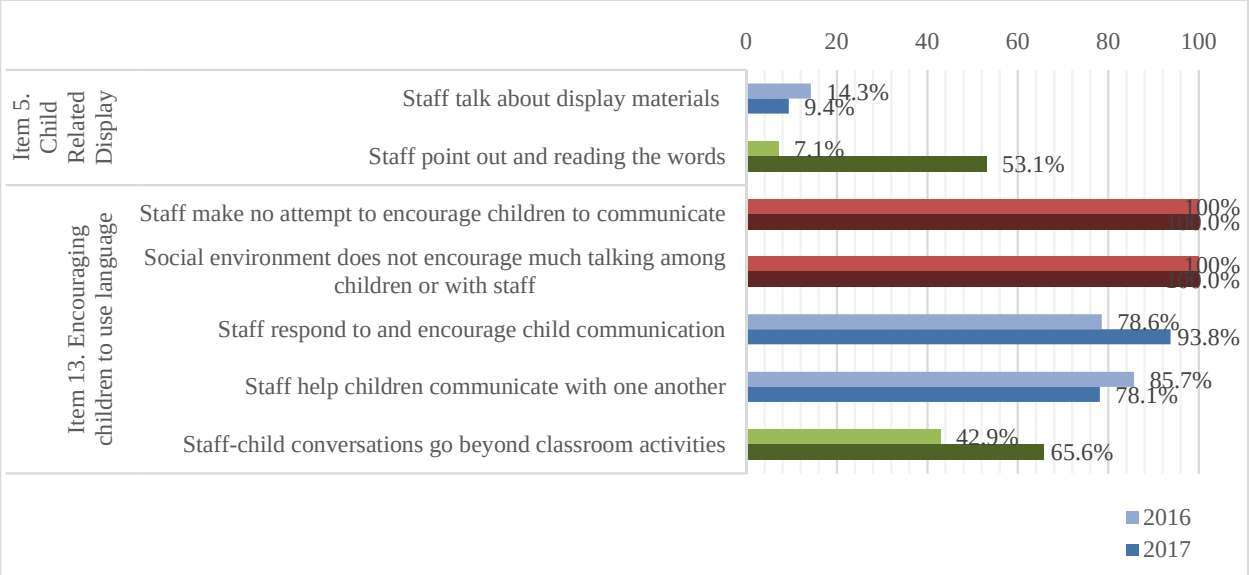
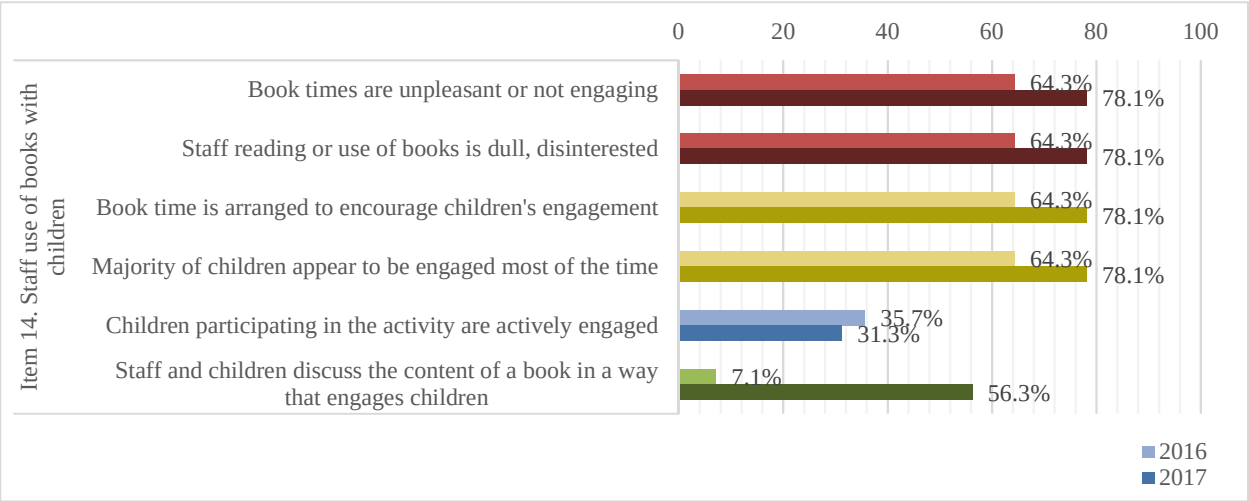


Figure 23 compares the indicators for use of books by staff across the two years. The majority of classrooms had increases in all but one of these indicators. In particular, there was a strong increase in the percentage of classrooms were staff and children were discussing the content of books in an engaging way (from 7% to 56%). This is an indicator of level 7 or "excellent'.

[11] Sometimes, not all items have indicators about engagement that range the full course (inadequate, minimal, good and excellent) so what is presented here is that which is captured by the tool on engagement regardless of the level.

Figure 23. Use of books (N=32 in 2017, N=14 in 2016)



Indicators of interest for Item 15 (Encouraging children's use of books), Item 16 (Print) and Item 17 (Fine Motor) are shown in Figure 24. There has been an increase in the percentage of classrooms meeting level 5 indicators (blue) and level 7 indicators (green) here as well. The largest increases are observed for books organized in a defined reading interest center (from 64% to 94%), books displayed in a way that encourages book use (34% to 75% of classrooms), and staff showing extended interest in what children create with materials (34% to 63%).

Figure 24. Books and fine motor (N=32 in 2017, N=14 in 2016)

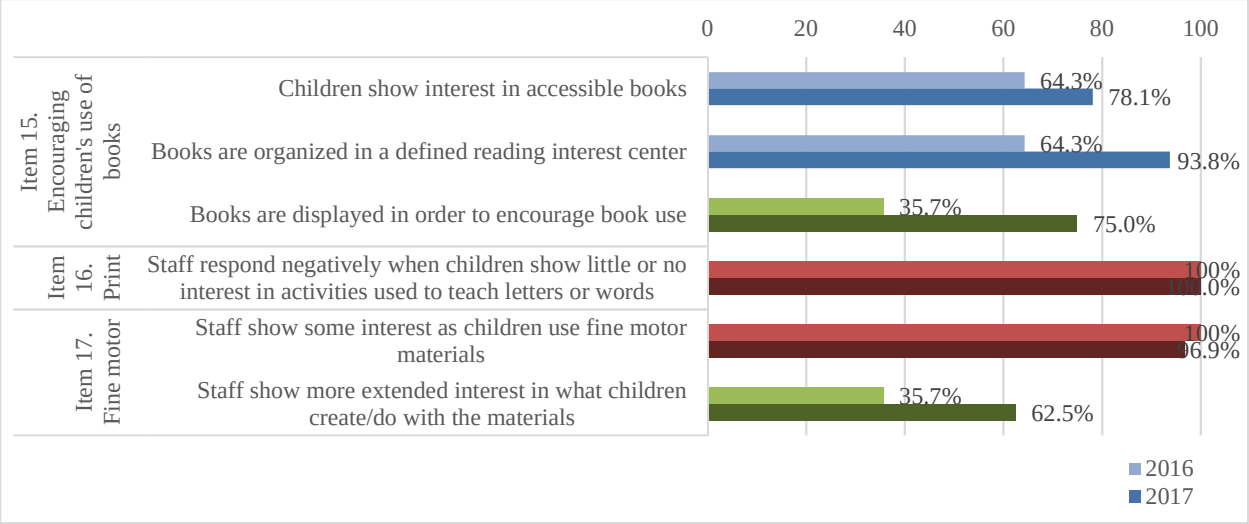


Figure 25 summarizes the indicators of interest in Art, Music and Movement and Blocks. Minimal changes are observed for these indicators. There is a slight increase in the percentage of classrooms in which staff have conversations with interested children about their work, and conversations with children about their block play, while simultaneously a decrease in the percentage of classrooms in which pleasant and engaged staff-led group music activities.

Figure 25. Art, music and blocks (N=32 in 2017, N=14 in 2016)

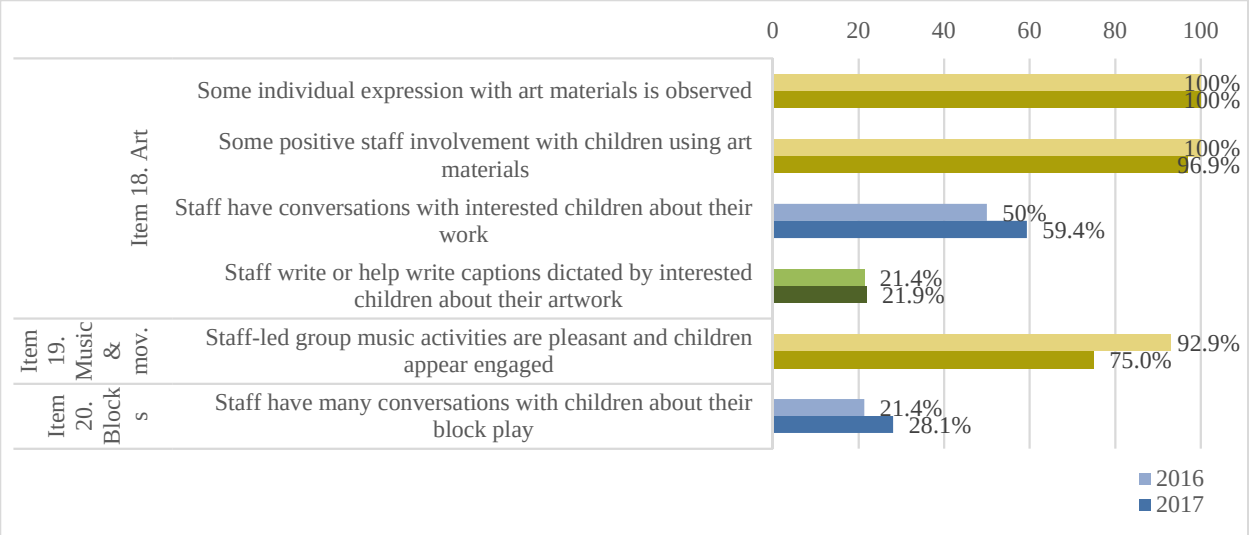
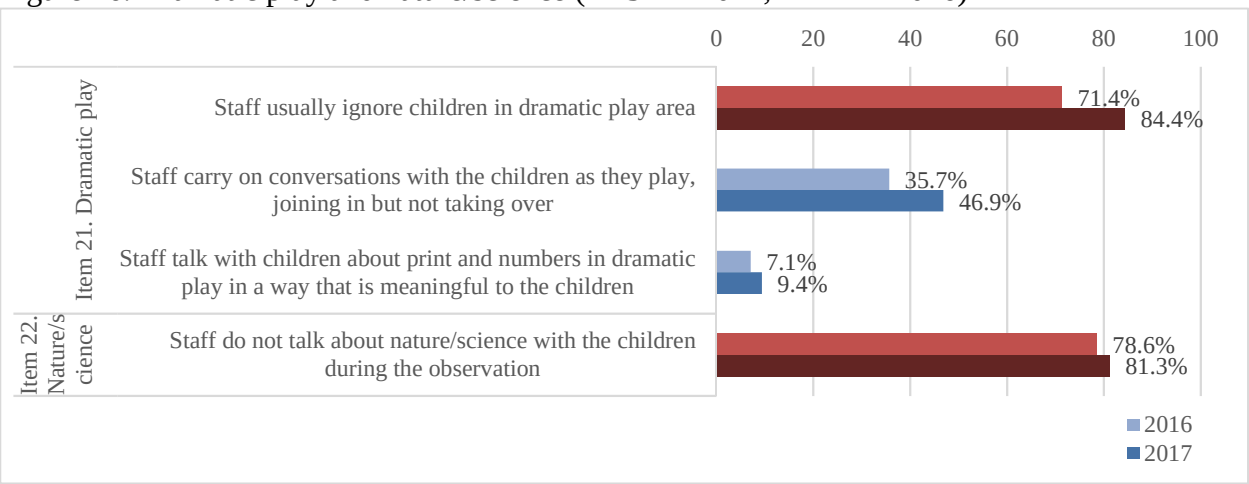


Figure 26 shows indicators for dramatic play and nature/science that relate to staff/child interactions. All four of these indicators had some increases in the percentage of classrooms where they were observed, including the level 5 ones (in blue).

Figure 26. Dramatic play and nature/science (N=32 in 2017, N=14 in 2016)



Indicators on interactions related to math and numbers are shown in Figure 27. Three of these shown improvements, with improvements observed in two level 5 items (blue): staff encourage math learning as part of daily routines (21% to 31%) and staff engage in conversations about math as they play in non-math areas (14% to 38%).

Figure 27. Math and numbers (N=32 in 2017, N=14 in 2016)

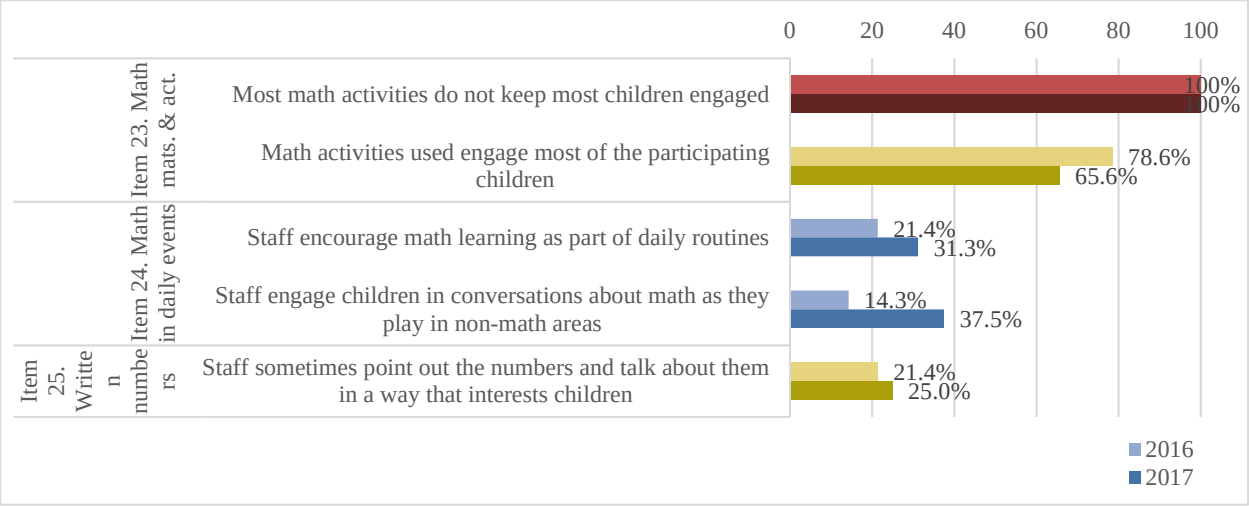
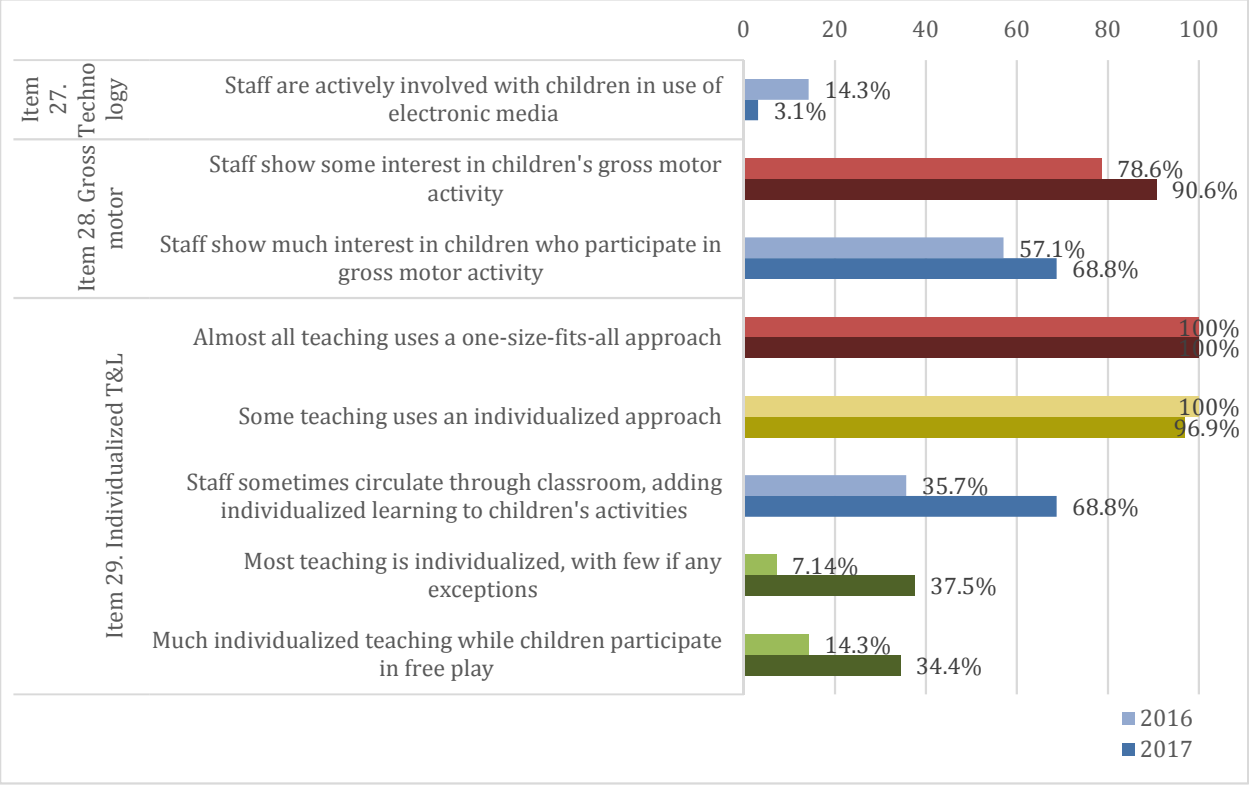


Figure 28 summarizes indicators relating to technology, gross motor activities and individualized teaching and learning. There were increases in indicators on interest in gross motor activities and staff circulating about the classroom adding to children's individualized learning (level 5 indicators) and on teaching being individualized and while children are in free play (level 7 indicators). These two increased from 7% to 38% and 14% to 34%, respectively.

Figure 28. Technology, gross motor and individualized teaching and learning (N=32 in 2017, N=14 in 2016)



In sum, the fact that the quality of interactions increased and better interactions between children and staff are observed in a higher percentage of classrooms while the program went from 14 classrooms in 2015–16 to 32 classrooms in 2016–17 is an important finding. Across all 47 indicators of interactions with children, the SPP program improved on 70%.

References

Aikens, N., Klein, A. K., Tarullo, L., & West, J. (2013). Getting ready for kindergarten: Children's progress during Head Start. FACES 2009 Report. (OPRE Report 2013–21a) Office of Planning. Research and Evaluation, Administration for Children and Families. Washington, D.C.: US Department of Health and Human Services.

Barnett, W.S. (2013). Expanding Access to Quality Pre-K is Sound Public Policy. New Brunswick, NJ: National Institute for Early Education Research.

Blair, C., & Razza, R. P. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. Child development, 78(2), 647-663.

Burchinal, M., Vandergrift, N., Pianta, R., & Mashburn, A. (2010). Threshold analysis of association between child care quality and child outcomes for low-income children in pre-kindergarten programs. Early Childhood Research Quarterly, 25(2), 166-176.

Childcare Quality & Early Learning Center for Research & Professional Development (Unpublished). Early Achievers Standards Validation Study. Seattle: University of Washington

Childcare Quality & Early Learning Center for Research & Professional Development (Unpublished). Large Scale Psychometric Assessment of ECERS 3. Seattle: University of Washington

Diamond, A., & Taylor, C. (1996). Development of an aspect of executive control: Development of the abilities to remember what I said and to "Do as I say, not as I do". Developmental psychobiology, 29(4), 315-334.

Dunn, L. M., & Dunn, D. M. (2007). PPVT-4: Peabody picture vocabulary test. Pearson Assessments.

Early, D. M., Maxwell, K. L., Burchinal, M., Alva, S., Bender, R. H., Bryant, D. Cai, K., Clifford, R.M., Ebanks, C., Griffin, J.A & Henry, G. T. (2007). Teachers' education, classroom quality, and young children's academic skills: Results from seven studies of preschool programs. Child development, 78(2), 558-580.

Edvance Research (2014). Pre-K 4 SA Evaluation Report. YEAR 1. Final Report Submitted to Early Childhood Education Municipal Development Corporation. San Antonio, TX: Author.

Harms, T., Clifford, R. M., & Cryer, D. (2014). Early childhood environment rating scale. Teachers College Press.

Hatfield, B. E., Burchinal, M. R., Pianta, R. C., & Sideris, J. (2016). Thresholds in the association between quality of teacher–child interactions and preschool children's school readiness skills. Early Childhood Research Quarterly, 36, 561-571.

Jenson, D. (2015) ECERS-3: One year out. Four States' Experiences with Planning and Implementing Use of ECERS-3. Presented at the QRIS National Meeting 2015. Available at http://www.qrisnetwork.org/sites/all/files/conference-session/resources/703ECERS3_0.pdf

Lamy, C.E., Frede, E. Seplocha, H., Strasser, J., Jambunathan, S., Juncker, J. A., Ferrar, H. Wiley, L., & Wolock, E. (2004). Inch by Inch, Row by Row Gonna Make This Garden Grow. Classroom Quality and Language Skills in the Abbott Preschool Program. Year One Report, 2002-2003 Early Learning Improvement Consortium. New Jersey: Early

Learning Improvement Consortium. Available at http://www.state.nj.us/education/ece/research/inch.pdf

Meador, D. N., Turner, K. A., Lipsey, M. W., & Farran, D. C. (2013). Administering Measures from the PRI Learning-Related Cognitive Self- Regulation Study. Nashville, TN: Peabody Research Institute. Available at https://my.vanderbilt.edu/cogselfregulation/files/2012/11/SR-Measure-Training-Manual-final.pdf

NIEER (2014). New Jersey Abbott Preschool Quality Evaluation Study. Summary Report. New Brunswick: Author.

NIEER (2016). New Jersey Abbott Preschool Quality Evaluation Study. Summary Report. New Brunswick: Author.

NYC Department of Education (2017). Pre-K Program Assessments Early Childhood Environmental Rating Scale –Revised (ECERS-R) and Classroom Assessment Scoring System (CLASS) Release. New York: Author. Available at http://schools.nyc.gov/NR/rdonlyres/5FEA3D5B-E615-4E16-83A8-C4E58A4D6F02/0/201516ProgramAssessmentResultsSummary.pdf

NYC Department of Education. (2015). Pre-K Program Assessments Classroom Assessment Scoring System (CLASS) and Early Childhood Environmental Rating Scale – Revised (ECERS-R) Release. New York: Author. Available at http://schools.nyc.gov/NR/rdonlyres/A8A27BFE-7C58-4F03-8EB7-B90E01BA3D0D/0/CLASSandECERSRReleaseDeckFinal.pdf

NYC Department of Education. (December 18, 2015). Mayor de Blasio Announces Over 68,500 Students Enrolled in Pre-K for All. Press office. New York: Author. Available at http://www1.nyc.gov/office-of-the-mayor/news/954-15/mayor-de-blasio-over-68-500-students-enrolled-pre-k-all

Office of Head Start. U.S. A National Overview of Grantee CLASS® Scores in 2015. Washington, D.C.: Department of Health and Human Services. Available at http://eclkc.ohs.acf.hhs.gov/hslc/data/class-reports/docs/national-class-2015-data.pdf

PAKEYS (Unpublished). What does the data tell us? The evolution of environment rating scale (ERS) use within QRIS. Available at https://qrisnetwork.org/sites/all/files/conference-session/resources/651DataTellsUs.pdf.

Phillips, D. A., Gormley, W. T., & Lowenstein, A. E. (2009). Inside the pre-kindergarten door: Classroom climate and instructional time allocation in Tulsa's pre-K programs. Early Childhood Research Quarterly, 24(3), 213-228.

Pianta, R. C., La Paro, K. M., & Hamre, B. K. (2008). Classroom Assessment Scoring System: Manual Pre-K. Education Review//Reseñas Educativas.

Qi, C. H., Kaiser, A. P., Milan, S., & Hancock, T. (2006). Language performance of low-income African American and European American preschool children on the PPVT–III. Language, Speech, and Hearing Services in Schools, 37(1), 5-16.

Rivers, N. M. (2016). Seattle Public Schools and Housing Report. Seattle: Seattle Public Schools. Available at https://www.seattleschools.org/cms/one.aspx?portalId=627&pageId=15652.

Weiland, C., Ulvestad, K., Sachs, J., & Yoshikawa, H. (2013). Associations between classroom quality and children's vocabulary and executive function skills in an urban public prekindergarten program. Early Childhood Research Quarterly, 28(2), 199-209.

Wong, V. C., Cook, T. D., Barnett, W. S., & Jung, K. (2008). An effectiveness-based evaluation of five state pre-kindergarten programs. Journal of policy Analysis and management, 27(1), 122-154.

Woodcock, R. W., McGrew, K. S., Mather, N., & Schrank, F. (2001). Woodcock-Johnson III NU tests of achievement. Rolling Meadows, IL: Riverside Publishing.

Zelazo, P. D. (2006). The dimensional change card sort (DCCS): A method of assessing executive function in children. Nature Protocols, 1, 297-301.

Appendices

Appendix A. ECERS-3 and CLASS scores by Agency and Class Size. Item level.
Appendix B.1. Raw Score Tables.
Appendix B.2. Standard and Raw Score Tables DEEL Demographics.
Appendix C.1. Sensitivity Analyses.
Appendix C.2. Sensitivity Analyses with DEEL Demographics.
Appendix D. Analyses for moderators.
Appendix E. Indicators tables for interactions with children.
Appendix F. Family Survey.

Appendix A. ECERS-3 and CLASS scores by Class Size and Item level.

Table A.1. ECERS-3 Item, Subscale, and Overall Means by Class Size, 2016 & 2017

ECERS-3 Item and Subscales	Spring 2016		Spring 2017	
	Small (≤18) (N=6)	Large (>18) (N=8)	Small (≤18) (N=24)	Large (>18) (N=8)
Overall	**3.52**	**3.60**	**3.90**	**3.87**
Space and Furnishings	3.91	3.86	3.89	4.07
1. Indoor space	6.33	6.50	5.42	5.63
2. Furnishings for care, play and learning	4.50	4.25	4.38	5.13
3. Room arrangement for play and learning	3.67	3.63	4.79	4.50
4. Space for privacy	4.17	4.13	4.38	5.00
5. Child-related display	3.33	3.38	2.92	3.63
6. Space for gross motor play	3.50	2.88	3.17	2.75
7. Gross motor equipment	1.83	2.25	2.21	1.88
Personal Care Routines	3.33	3.00	3.30	3.72
8. Meals/ snacks	3.33	2.88	3.79	4.13
9. Toileting/diapering	2.33	2.13	3.08	3.50
10. Health practices	3.17	2.75	2.58	3.00
11. Safety practices	4.50	4.25	3.75	4.25
Language and Literacy	3.23	3.65	3.87	4.10
12. Helping children expand vocabulary	3.33	3.63	3.58	3.75
13. Encouraging children to use language	4.00	4.63	4.79	5.00
14. Staff use of books with children	3.17	3.00	3.46	3.63
15. Encouraging children's use of books	4.17	4.25	4.42	4.38
16. Becoming familiar with print	1.50	2.75	3.08	3.75
Learning Activities	2.61	3.06	3.26	3.28
17. Fine motor	4.17	4.50	4.50	4.38
18. Art	3.50	3.88	4.17	4.63
19. Music and movement	3.33	3.63	3.38	3.75
20. Blocks	1.33	2.50	3.04	2.75
21. Dramatic Play	2.67	2.88	3.63	3.13
22. Nature/science	2.33	2.63	2.25	2.38
23. Math materials and activities	1.50	1.88	2.04	2.88
24. Math in daily events	2.50	3.13	3.33	3.38
25. Understanding written numbers	1.00	1.50	1.71	1.63
26. Promoting acceptance of diversity	4.00	4.38	4.54	3.75
Interaction	4.47	4.50	5.21	4.35
27. Appropriate use of technology	1.00	1.00	3.00	4.00
28. Supervision of gross motor	4.00	3.50	5.04	3.13
29. Individualized teaching and learning	3.83	4.50	5.00	4.75
30. Staff-child interaction	5.50	4.50	5.96	4.75
31. Peer interaction	4.83	5.13	4.92	4.63
32. Discipline	4.17	4.88	5.13	4.50
Program Structure	4.89	4.09	4.76	4.38
33. Transitions and waiting times	5.50	4.38	5.00	4.00
34. Free play	4.83	4.25	4.42	4.50
35. Whole -group activities for play and learning	4.33	3.63	4.88	4.63

Table A.2. CLASS Dimension and Domain Means by Class Size, 2016 & 2017

CLASS Dimensions and Domains	Spring 2016		Spring 2017	
	Small (≤18) (N=6)	Large (>18) (N=8)	Small (≤18) (N=24)	Large (>18) (N=8)
Emotional Support Domain	6.08	6.18	6.27	6.37
1. Positive Climate	5.67	5.91	6.30	6.41
2. Negative Climate*	1.04	1.22	1.07	1.00
3. Teacher Sensitivity	5.79	6.00	6.08	5.91
4. Regard for Student Perspectives	5.88	6.03	5.99	5.88
Classroom Organization Domain	5.60	5.72	5.54	5.56
5. Behavior Management	5.71	5.75	5.41	5.63
6. Productivity	5.92	6.16	5.82	6.16
7. Instructional Learning Formats	5.17	5.25	5.20	5.25
Instructional Support Domain	2.43	2.82	3.13	2.85
8. Concept Development	1.88	2.22	2.73	2.38
9. Quality of Feedback	2.29	2.84	3.09	2.84
10. Language Modeling	3.13	3.41	3.52	3.72

*The Negative Climate dimension was transposed so that on here, high represents "good"

Table A.3. ECERS and CLASS Dimension and Domain Means by Child Demographics, 2016 & 2017 (N=291)

		N	ECERS		CLASS_ES		CLASS_CO		CLASS_IS	
			Mean	SD	Mean	SD	Mean	SD	Mean	SD
Total		291	3.89	0.57	6.30	0.48	5.52	0.73	2.99	0.81
Gender	Female	141	3.87	0.60	6.26	0.51	5.51	0.72	3.01	0.79
	Male	150	3.91	0.54	6.35	0.45	5.53	0.75	2.98	0.82
Age	3-Year-Olds	47	3.82	0.45	6.17	0.47	5.20	0.75	2.68	0.59
	4-Year-Olds	244	3.90	0.59	6.33	0.47	5.58	0.71	3.05	0.83
Ethnicity	White	62	3.94	0.68	6.34	0.48	5.67	0.65	3.12	0.81
	Black	67	3.90	0.41	6.34	0.39	5.34	0.80	2.91	0.86
	Asian	48	3.92	0.60	6.26	0.46	5.63	0.64	3.15	0.75
	Hispanic	23	3.73	0.58	6.35	0.49	5.43	0.82	3.02	0.84
	Other	91	3.86	0.58	6.27	0.54	5.52	0.74	2.88	0.77
Language	English	199	3.88	0.59	6.30	0.48	5.53	0.70	2.96	0.79
	Bilingual	46	3.98	0.60	6.35	0.41	5.63	0.72	3.18	0.69
	Unknown	46	3.84	0.46	6.29	0.54	5.38	0.85	2.96	0.96
FPL	<100	61	3.92	0.48	6.31	0.39	5.36	0.64	2.89	0.68
	100-300	137	3.89	0.54	6.27	0.50	5.45	0.83	2.96	0.81
	>300	93	3.86	0.67	6.35	0.50	5.73	0.59	3.11	0.86

Appendix B.1. Raw Score Tables.

Table B.1.1. Receptive vocabulary raw score means and gains by child characteristics

		N	PPVT Raw 2016 Fall		PPVT Raw 2017 Spring		PPVT Raw Gains	
			Mean	SD	Mean	SD	Mean	SD
Total		288	73.34	27.14	85.52	25.97	12.18	12.77
Gender	Female	139	74.46	26.49	87.65	24.61	13.19	12.08
	Male	149	72.29	27.79	83.53	27.11	11.24	13.35
Age	3-Year Olds	46	45.3	18.18	60.52	17.02	15.22	12.00
	4-Year Olds	242	78.67	25.24	90.27	24.64	11.60	12.85
Ethnicity	White	61	97.54	20.09	108.64	17.47	11.10	12.35
	Black	66	59.12	23.48	69.82	23.04	10.70	13.16
	Asian	48	61.42	25.24	76.02	24.56	14.60	10.92
	Hispanic	22	57.64	24.16	68.00	20.15	10.36	12.47
	Other	91	77.51	22.69	90.65	21.48	13.14	13.68
Language	English	198	81.22	24.76	92.72	24.33	11.50	13.59
	Spanish	3	59.00	20.95	69.33	23.67	10.33	7.51
	Vietnamese	12	50.17	19.91	63.5	22.05	13.33	9.30
	Other	31	57.13	23.84	71.74	23.43	14.61	9.02
	Unknown	44	56.57	25.75	69.91	21.85	13.34	12.33
FPL	<100	59	59.39	24.46	71.80	24.87	12.41	10.62
	100-300	93	86.62	26.5	98.12	25.07	11.49	12.60
	>300	136	70.30	24.8	82.85	23.26	12.55	13.76

Table B.1.2. Receptive vocabulary raw score means and gains by center characteristics

			PPVT Raw 2016 Fall		PPVT Raw 2017 Spring		PPVT Raw Gains	
			Mean	SD	Mean	SD	Mean	SD
Total		288	73.34	27.14	85.52	25.97	12.18	12.77
Agency	Agency 1	20	62.00	29.41	75.70	24.96	13.70	9.66
	Agency 2	10	48.3	13.64	59.90	21.66	11.60	11.06
	Agency 3	14	93.43	27.90	105.21	24.55	11.79	12.69
	Agency 4	61	74.74	25.16	87.07	25.03	12.33	11.92
	Agency 5	11	52.91	21.63	59.09	17.62	6.18	12.32
	Agency 6	95	74.77	24.29	88.17	23.63	13.40	12.30
	Agency 7	23	76.70	25.02	85.13	18.26	8.43	13.83
	Agency 8	28	74.00	31.93	89.46	29.78	15.46	12.94
	Agency 9	8	72.25	29.55	86.88	26.93	14.63	10.90
	Agency 10	18	79.56	31.94	86.00	31.86	6.44	18.99
Class Size	18 or Less	187	70.50	27.64	82.12	26.44	11.61	13.01
	More than 18	101	78.58	25.51	91.81	23.97	13.23	12.30
Curriculum	Creative Curriculum	84	69.11	30.49	81.68	29.05	12.57	13.59
	HighScope	204	75.08	25.52	87.1	24.50	12.02	12.44
ECERS	Less than 3	14	65.36	25.21	82.86	26.17	17.50	11.40
	3 or More	274	73.74	27.22	85.65	26.00	11.91	12.79
CLASS ES	Less than 5.5	16	67.94	27.93	86.06	27.93	18.13	12.48
	5.5 or More	272	73.65	27.12	85.49	25.91	11.83	12.72
CLASS CO	Less than 5.5	106	68.83	26.94	80.91	26.90	12.08	13.44
	5.5 or More	182	75.96	26.99	88.20	25.10	12.24	12.39
CLASS IS	Less than 3	156	70.54	28.02	82.25	26.44	11.71	13.85
	3 or More	132	76.64	25.79	89.38	24.96	12.73	11.37

Table B.1.3. Literacy raw score means and gains by child characteristics

		N	WJ-LW Raw 2016 Fall		WJ-LW Raw 2017 Spring		WJ-LW Raw Gains	
			Mean	SD	Mean	SD	Mean	SD
Total		286	8.36	6.35	11.23	7.22	2.88	3.31
Gender	Female	138	8.13	6.45	11.21	7.49	3.08	3.54
	Male	148	8.57	6.28	11.26	7.00	2.69	3.08
Age	3-Year-Old Cohort	47	4.83	3.89	6.96	4.33	2.13	3.16
	4-Year-Old Cohort	239	9.05	6.52	12.08	7.39	3.03	3.32
Ethnicity	White	61	10.75	6.81	13.93	7.43	3.18	2.82
	Black	65	7.35	6.07	10.34	7.31	2.98	3.55
	Asian	48	9.27	6.70	12.81	8.36	3.54	4.26
	Hispanic	21	4.52	3.61	6.43	3.61	1.90	2.47
	Other	91	7.87	5.97	10.34	6.17	2.47	2.98
Language	English	196	8.72	6.52	11.43	7.13	2.71	3.01
	Spanish	3	7.33	5.51	8.67	6.51	1.33	2.52
	Vietnamese	12	7.83	4.97	9.75	4.96	1.92	2.94
	Other	31	9.06	7.23	13.39	9.52	4.32	4.61
	Unknown	44	6.43	5.05	9.41	5.99	2.98	3.46
FPL	English	196	8.72	6.52	11.43	7.13	2.71	3.01
	Bilingual	46	8.63	6.53	12.13	8.47	3.50	4.24
	Unknown	44	6.43	5.05	9.41	5.99	2.98	3.46

Table B.1.4. Literacy raw score means and gains by center characteristics

			WJ-LW Raw 2016 Fall		WJ-LW Raw 2017 Spring		WJ-LW Raw Gains	
			Mean	SD	Mean	SD	Mean	SD
Total		286	8.36	6.35	11.23	7.22	2.88	3.31
Agency	Agency 1	19	5.63	4.02	9.00	6.27	3.37	3.17
	Agency 2	10	11.7	10.37	15.4	12.84	3.70	4.32
	Agency 3	14	6.21	3.77	9.86	5.26	3.64	2.59
	Agency 4	61	8.39	4.91	10.57	5.59	2.18	3.36
	Agency 5	11	7.09	4.81	8.73	5.12	1.64	2.11
	Agency 6	93	9.03	7.58	12.56	8.48	3.53	3.29
	Agency 7	23	8.61	4.39	10.22	4.27	1.61	2.17
	Agency 8	29	8.45	6.03	11.76	7.06	3.31	3.57
	Agency 9	8	7.00	4.69	10.63	4.69	3.63	4.03
	Agency 10	18	8.33	8.10	10.00	7.69	1.67	3.53
Class Size	18 or Less	187	8.27	6.59	10.90	7.18	2.63	3.17
	More than 18	99	8.53	5.91	11.87	7.30	3.34	3.53
Curriculum	Creative Curriculum	84	8.04	6.77	11.08	7.80	3.05	3.6
	HighScope	202	8.49	6.18	11.3	6.99	2.81	3.19
ECERS	Less than 3	15	7.27	5.18	9.00	4.55	1.73	2.6
	3 or More	271	8.42	6.41	11.36	7.33	2.94	3.34
CLASS ES	Less than 5.5	17	6.71	4.91	10.00	6.91	3.29	4.19
	5.5 or More	269	8.46	6.43	11.31	7.25	2.85	3.25
CLASS CO	Less than 5.5	103	7.14	5.39	9.60	5.90	2.47	3.35
	5.5 or More	183	9.04	6.75	12.15	7.74	3.11	3.27
CLASS IS	Less than 3	154	8.08	6.22	10.86	7.11	2.78	3.45
	3 or More	132	8.67	6.52	11.67	7.36	2.99	3.15

Table B.1.5. Math raw score means and gains by child characteristics

		N	WJ-AP Raw 2016 Fall		WJ-AP Raw 2017 Spring		WJ-AP Raw Gains	
			Mean	SD	Mean	SD	Mean	SD
Total		286	10.99	5.58	13.93	5.32	2.94	3.53
Gender	Female	138	10.83	5.72	13.37	5.25	2.54	3.68
	Male	148	11.15	5.46	14.45	5.35	3.30	3.36
Age	3-Year-Old Cohort	47	4.79	3.91	8.64	4.11	3.85	4.43
	4-Year-Old Cohort	239	12.21	5.02	14.97	4.90	2.76	3.31
Ethnicity	White	61	15.26	4.36	16.95	4.27	1.69	3.64
	Black	65	7.89	5.27	10.94	5.28	3.05	3.17
	Asian	48	10.08	5.70	13.83	5.42	3.75	3.97
	Hispanic	21	7.05	4.15	11.00	4.31	3.95	2.73
	Other	91	11.74	4.63	14.77	4.77	3.03	3.47
Language	English	196	12.35	4.91	14.92	4.99	2.57	3.63
	Spanish	3	6.67	5.51	11.33	5.51	4.67	0.58
	Vietnamese	12	9.00	5.48	13.42	6.08	4.42	3.8
	Other	31	9.29	6.19	13.00	5.59	3.71	3.28
	Unknown	44	6.98	5.63	10.50	4.94	3.52	3.11
FPL	English	196	12.35	4.91	14.92	4.99	2.57	3.63
	Bilingual	46	9.04	5.89	13.00	5.61	3.96	3.29
	Unknown	44	6.98	5.63	10.50	4.94	3.52	3.11

Table B.1.6 Math raw score means and gains by center characteristics

			WJ-AP Raw 2016 Fall		WJ-AP Raw 2017 Spring		WJ-AP Raw Gains	
			Mean	SD	Mean	SD	Mean	SD
Total		286	10.99	5.58	13.93	5.32	2.94	3.53
Agency	Agency 1	19	7.26	6.51	11.47	5.10	4.21	4.05
	Agency 2	10	6.00	4.57	8.10	4.86	2.10	4.61
	Agency 3	14	14.14	5.27	14.71	4.05	0.57	2.77
	Agency 4	61	11.79	4.87	13.84	5.13	2.05	4.02
	Agency 5	11	8.27	4.96	14.64	3.80	6.36	3.14
	Agency 6	93	11.82	5.3	14.91	5.39	3.10	3.38
	Agency 7	23	10.04	4.93	13.48	4.76	3.43	2.37
	Agency 8	29	10.72	5.88	13.69	5.75	2.97	3.25
	Agency 9	8	12.38	4.93	15.75	3.77	3.38	2.83
	Agency 10	18	11.00	6.66	14.11	6.20	3.11	2.14
Class Size	18 or Less	187	10.04	5.79	13.18	5.48	3.14	3.36
	More than 18	99	12.8	4.66	15.34	4.70	2.55	3.82
Curriculum	Creative Curriculum	84	9.60	6.22	12.81	5.74	3.21	3.38
	HighScope	202	11.57	5.20	14.4	5.07	2.82	3.6
ECERS	Less than 3	15	8.40	5.44	11.27	5.54	2.87	2.36
	3 or More	271	11.14	5.56	14.08	5.28	2.94	3.59
CLASS ES	Less than 5.5	17	8.41	5.71	11.65	6.08	3.24	2.36
	5.5 or More	269	11.16	5.54	14.07	5.25	2.92	3.6
CLASS CO	Less than 5.5	103	9.90	5.68	12.7	5.68	2.80	3.69
	5.5 or More	183	11.61	5.44	14.62	4.99	3.02	3.45
CLASS IS	Less than 3	154	10.36	5.51	13.51	5.56	3.15	3.65
	3 or More	132	11.73	5.58	14.42	5.00	2.69	3.39

Appendix B.2. Standard and Raw Score Tables DEEL Demographics.

Table B.2.1. Standard score means and gains by child characteristics

Race/Ethnicity		N	2016 Fall		2017 Spring		Gains	
			Mean	SD	Mean	SD	Mean	SD
PPVT	White	64	115.51	13.77	116.95	12.11	1.44	9.72
Standard	Black	61	90.18	14.66	91.87	14.29	1.69	10.02
	Asian	48	91.54	16.76	96.21	15.51	4.67	8.46
	Hispanic	41	98.73	17.18	102.71	15.96	3.98	8.78
	Other	71	103.97	15.27	106.00	14.56	2.03	10.34
WJ-LW	White	64	106.08	14.73	106.10	14.41	0.02	6.83
Standard	Black	61	98.66	16.17	101.28	16.46	2.62	10.05
	Asian	48	104.25	16.77	107.06	16.82	2.81	11.81
	Hispanic	41	94.76	14.01	94.83	12.34	0.07	10.66
	Other	71	100.25	15.06	99.25	14.01	-1.00	8.49
WJ-AP	White	64	113.57	9.99	110.29	10.47	-3.29	11.75
Standard	Black	61	94.11	14.34	98.56	11.15	4.44	12.82
	Asian	48	100.27	15.78	105.13	14.72	4.85	13.80
	Hispanic	41	98.15	14.45	102.41	12.47	4.27	11.03
	Other	71	104.30	11.48	105.97	12.92	1.68	10.21

Table B.2.2. Standard score means and gains by child characteristics

Race/Ethnicity		N	2016 Fall		2017 Spring		Gains	
			Mean	SD	Mean	SD	Mean	SD
PPVT	White	64	96.16	20.51	107.17	17.84	11.02	12.93
Raw	Black	61	57.39	22.83	68.05	22.78	10.66	13.53
	Asian	48	60.08	23.86	74.58	23.56	14.50	10.70
	Hispanic	41	69.83	24.96	83.85	24.05	14.02	11.87
	Other	71	77.79	24.17	89.87	22.53	12.08	13.97
WJ-LW	White	64	106.08	14.73	13.83	7.26	3.14	2.80
Raw	Black	61	98.66	16.17	10.26	7.47	2.90	3.38
	Asian	48	104.25	16.77	13.10	8.18	3.85	4.22
	Hispanic	41	94.76	14.01	8.05	5.22	1.93	2.74
	Other	71	100.25	15.06	10.37	6.43	2.30	2.86
WJ-AP	White	64	15.41	3.88	16.98	4.05	1.57	3.57
Raw	Black	61	8.00	5.37	10.82	5.33	2.82	3.20
	Asian	48	9.94	5.67	13.81	5.35	3.88	3.96
	Hispanic	41	9.34	5.19	12.95	5.07	3.61	3.26
	Other	71	11.62	4.48	14.76	4.74	3.14	3.36
DCCS	White	64	1.78	0.49	2.03	0.60	0.25	0.60
	Black	61	1.26	0.71	1.51	0.72	0.25	0.70
	Asian	48	1.52	0.58	1.65	0.57	0.13	0.53
	Hispanic	41	1.39	0.63	1.61	0.63	0.22	0.48
	Other	71	1.54	0.69	1.75	0.67	0.21	0.63
PT	White	64	9.48	4.83	10.79	5.41	1.32	5.20
	Black	61	3.84	5.25	6.39	5.82	2.56	5.30
	Asian	48	6.50	6.53	9.10	6.15	2.60	6.27
	Hispanic	41	5.27	5.08	7.39	6.19	2.12	4.92
	Other	71	6.30	5.73	9.24	5.99	2.94	5.14

Table B.2.3. Standard score means and gains by child characteristics

Language		N	2016 Fall		2017 Spring		Gains	
			Mean	SD	Mean	SD	Mean	SD
PPVT	English	195	105.79	16.59	107.70	16.25	1.91	10.14
Standard	Spanish	3	89.33	18.04	91.33	17.21	2.00	6.25
	Vietnamese	12	85.25	15.51	88.92	15.01	3.67	7.79
	Other	31	89.23	14.76	93.97	14.43	4.74	6.60
	Unknown	43	91.05	16.81	94.67	13.09	3.63	9.71
WJ-LW	English	195	101.52	15.73	101.61	15.22	0.09	8.74
Standard	Spanish	3	96.00	21.00	95.33	17.79	-0.67	6.35
	Vietnamese	12	102.25	13.61	100.67	13.54	-1.58	10.09
	Other	31	103.84	17.47	108.16	18.61	4.32	8.84
	Unknown	43	97.14	14.70	99.42	13.46	2.28	12.71
WJ-AP	English	195	105.88	12.92	106.46	12.67	0.58	11.75
Standard	Spanish	3	90.00	19.52	98.33	17.39	8.33	5.51
	Vietnamese	12	97.50	17.48	104.08	18.68	6.58	15.65
	Other	31	98.29	16.44	102.84	12.73	4.55	9.55
	Unknown	43	93.14	14.76	98.53	9.79	5.40	14.27

Table B.2.4. Standard score means and gains by child characteristics

Language		N	2016 Fall		2017 Spring		Gains	
			Mean	SD	Mean	SD	Mean	SD
PPVT	English	198	81.22	24.76	92.72	24.33	11.50	13.59
Raw	Spanish	3	59.00	20.95	69.33	23.67	10.33	7.51
	Vietnamese	12	50.17	19.91	63.5	22.05	13.33	9.30
	Other	31	57.13	23.84	71.74	23.43	14.61	9.02
	Unknown	44	56.57	25.75	69.91	21.85	13.34	12.33
WJ-LW	English	196	8.72	6.52	11.43	7.13	2.71	3.01
Raw	Spanish	3	7.33	5.51	8.67	6.51	1.33	2.52
	Vietnamese	12	7.83	4.97	9.75	4.96	1.92	2.94
	Other	31	9.06	7.23	13.39	9.52	4.32	4.61
	Unknown	44	6.43	5.05	9.41	5.99	2.98	3.46
WJ-AP	English	196	12.35	4.91	14.92	4.99	2.57	3.63
Raw	Spanish	3	6.67	5.51	11.33	5.51	4.67	0.58
	Vietnamese	12	9.00	5.48	13.42	6.08	4.42	3.8
	Other	31	9.29	6.19	13.00	5.59	3.71	3.28
	Unknown	44	6.98	5.63	10.5	4.94	3.52	3.11
DCCS	English	195	1.59	0.646	1.82	0.661	0.23	0.601
	Spanish	3	1.67	0.577	1.67	0.577	0.00	0.00
	Vietnamese	12	1.50	0.522	1.58	0.669	0.08	0.515
	Other	31	1.35	0.755	1.61	0.615	0.26	0.682
	Unknown	44	1.20	0.509	1.39	0.618	0.18	0.582
PT	English	195	7.19	5.631	9.52	5.667	2.33	5.307
	Spanish	3	5.00	5.292	2.67	3.215	-2.33	2.517
	Vietnamese	12	3.92	5.728	8.58	7.513	4.67	6.344
	Other	31	5.48	5.938	7.32	6.063	1.84	6.17
	Unknown	44	3.89	5.723	6.25	6.588	2.36	4.636

Appendix C.1. Sensitivity Analyses.

Table C.1. Multivariate analyses of children's 2016–17 raw score gains in relation to child and site or classroom characteristics and ECERS-3

	Rec. Vocabulary Raw	Literacy Raw	Math Raw	DCCS Sum
Female	1.772	0.379	**-0.886***	0.556
	(1.32)	(0.37)	(0.36)	(0.52)
Black	**-6.465***	-0.665	-0.369	-0.356
	(2.59)	(0.69)	(0.70)	(0.97)
Asian	-2.008	0.594	0.244	-0.393
	(2.65)	(0.71)	(0.70)	(0.99)
Hispanic	**-8.109***	-1.739	0.066	-1.096
	(3.22)	(0.89)	(0.89)	(1.22)
Other Race	-0.350	-1.042	0.571	-1.072
	(2.03)	(0.56)	(0.55)	(0.77)
DLL	0.052	0.031	0.349	-0.290
	(2.28)	(0.63)	(0.61)	(0.88)
Agency Selected	**7.206***	-0.321	0.469	-1.489
	(3.40)	(0.95)	(0.92)	(1.33)
FPL <100%	1.132	**1.660***	-1.100	0.487
	(2.91)	(0.81)	(0.78)	(1.13)
FPL 100-300%	0.486	0.995	-0.361	0.246
	(2.27)	(0.64)	(0.62)	(0.89)
Income ≤20K	-4.549	-1.079	-0.828	-0.257
	(3.21)	(0.90)	(0.89)	(1.25)
Income 21K-40K	-2.222	-0.071	-0.572	0.375
	(2.84)	(0.79)	(0.77)	(1.11)
Income 41K-60K	1.892	-1.031	-0.437	0.546
	(2.89)	(0.81)	(0.79)	(1.14)
Income 61K-80K	**-5.200***	0.637	-0.732	0.433
	(2.40)	(0.67)	(0.65)	(0.94)
Class Size	0.423	-0.035	0.127	0.050
	(0.37)	(0.10)	(0.10)	(0.14)
ECERS	0.357	**1.374***	0.645	-0.446
	(1.95)	(0.54)	(0.53)	(0.76)
Observations	288	286	286	286

* p<0.05; ** p<0.01; *** p<0.001. Note: Reference groups omitted from the estimation are Males, White, English, FPL 300%+ and Income>80 thousand. Other controls are pre-test, age in months, days between tests, agencies, and an indicator for missing language or income. Raw scores are used for PPVT, and WJ or WM. Alternative scoring is used for the DCCS. Errors are clustered at the site level.

Table C.2. Multivariate analyses of children's 2016–17 raw score gains in relation to child and site or classroom characteristics and CLASS dimensions

	Rec. Vocabulary Raw	Literacy Raw	Math Raw	DCCS Sum
Female	1.528	0.301	**-0.902***	0.508
	(1.31)	(0.38)	(0.36)	(0.52)
Black	**-5.909***	-0.632	-0.383	-0.328
	(2.56)	(0.70)	(0.70)	(0.96)
Asian	-2.250	0.510	0.242	-0.445
	(2.63)	(0.72)	(0.70)	(0.99)
Hispanic	**-7.791***	-1.724	0.045	-0.997
	(3.18)	(0.90)	(0.89)	(1.21)
Other Race	-0.350	-1.040	0.572	-1.165
	(2.01)	(0.56)	(0.55)	(0.77)
Bilingual	-0.131	0.141	0.379	-0.209
	(2.25)	(0.63)	(0.61)	(0.87)
Agency Selected	**7.242***	-0.530	0.390	-1.276
	(3.38)	(0.97)	(0.93)	(1.33)
FPL <100%	2.189	**1.863***	-1.025	0.445
	(2.89)	(0.82)	(0.79)	(1.13)
FPL 100-300%	0.926	1.140	-0.269	0.146
	(2.25)	(0.64)	(0.62)	(0.89)
Income ≤20K	-5.108	-1.086	-0.856	-0.194
	(3.19)	(0.91)	(0.89)	(1.25)
Income 21K-40K	-2.475	-0.110	-0.610	0.458
	(2.81)	(0.80)	(0.78)	(1.11)
Income 41K-60K	1.757	-0.979	-0.452	0.585
	(2.87)	(0.82)	(0.80)	(1.14)
Income 61K-80K	**-5.286***	0.649	-0.744	0.522
	(2.38)	(0.68)	(0.65)	(0.93)
Class Size	0.360	-0.086	0.095	0.039
	(0.39)	(0.11)	(0.11)	(0.15)
CLASS_ES	-2.649	-0.199	0.274	-1.112
	(1.69)	(0.48)	(0.46)	(0.66)
CLASS_CO	2.587	0.444	0.154	0.487
	(1.61)	(0.47)	(0.45)	(0.65)
CLASS_IS	1.202	0.033	0.040	-0.544
	(1.09)	(0.32)	(0.31)	(0.44)
N	288	286	286	286

* p<0.05; ** p<0.01; *** p<0.001. Note: Reference groups omitted from the estimation are Males, White, English, FPL 300%+ and Income>80 thousand. Other controls are pre-test, age in months, days between tests, agencies, and an indicator for missing language or income. Raw scores are used for PPVT, and WJ or WM. Alternative scoring is used for the DCCS. Errors are clustered at the site level.

Table C.3. Multivariate analyses of children's 2016–17 standard score gains in relation to child and site or classroom characteristics and ECERS-3 thresholds.

				Executive Function	
	Rec. Vocabulary Standard	Literacy Standard	Math Standard	DCCS	PT
Female	1.248	0.264	-2.230	0.086	-0.575
	(0.98)	(1.05)	(1.16)	(0.06)	(0.54)
Black	**-4.536***	-0.653	-1.804	-0.020	-0.311
	(1.90)	(1.96)	(2.24)	(0.11)	(1.00)
Asian	-1.404	0.666	-0.842	-0.140	1.312
	(1.95)	(2.00)	(2.23)	(0.11)	(1.02)
Hispanic	**-5.660***	-4.524	-2.561	-0.179	-1.444
	(2.37)	(2.53)	(2.82)	(0.14)	(1.25)
Other Race	-0.490	**-3.151***	0.724	-0.151	-0.269
	(1.49)	(1.57)	(1.75)	(0.09)	(0.79)
Bilingual	0.193	0.299	0.753	-0.003	-0.777
	(1.67)	(1.77)	(1.95)	(0.10)	(0.90)
Agency Selected	4.617	0.778	2.648	-0.064	1.797
	(2.61)	(2.79)	(3.06)	(0.16)	(1.42)
FPL <100%	0.556	4.108	-4.003	-0.041	**-2.530***
	(2.15)	(2.28)	(2.51)	(0.13)	(1.16)
FPL 100-300%	0.085	1.888	-1.369	-0.061	-0.232
	(1.67)	(1.79)	(1.98)	(0.10)	(0.91)
Income ≤20K	-2.727	-1.535	-1.253	0.027	-1.517
	(2.37)	(2.52)	(2.82)	(0.14)	(1.29)
Income 21K-40K	-1.356	-0.440	-0.760	0.002	-1.496
	(2.09)	(2.24)	(2.47)	(0.13)	(1.14)
Income 41K-60K	1.930	-1.792	-1.077	0.079	-0.499
	(2.13)	(2.29)	(2.54)	(0.13)	(1.16)
Income 61K-80K	**-3.643***	1.905	-1.695	0.036	0.175
	(1.77)	(1.89)	(2.08)	(0.11)	(0.96)
Class Size	0.357	-0.181	0.289	0.011	0.209
	(0.27)	(0.29)	(0.32)	(0.02)	(0.15)
ECERS ≥3	-1.755	0.322	4.544	-0.066	-1.013
	(2.52)	(2.58)	(2.87)	(0.15)	(1.31)
N	288	286	286	286	285

* p<0.05; ** p<0.01; *** p<0.001. Note: Reference groups omitted from the estimation are Males, White, English, FPL 300%+ and Income>80 thousand. Other controls are pre-test, age in months, days between tests, agencies, and an indicator for missing language or income. Raw scores are used for PPVT, and WJ or WM. Errors are clustered at the site level.

Table C.4. Multivariate analyses of children's 2016–17 standard score gains in relation to child and site or classroom characteristics and CLASS dimension thresholds.

	Rec. Vocabulary	Literacy	Math	Executive Function	
	Standard	Standard	Standard	DCCS	PT
Female	1.225	0.117	**-2.338***	0.086	-0.548
	(0.97)	(1.04)	(1.16)	(0.06)	(0.54)
Black	**-4.468***	-0.524	-1.484	-0.033	-0.398
	(1.90)	(1.95)	(2.24)	(0.11)	(1.00)
Asian	-1.566	0.842	-0.305	-0.161	1.221
	(1.97)	(2.01)	(2.25)	(0.11)	(1.03)
Hispanic	**-5.699***	-4.286	-2.181	-0.189	-1.520
	(2.37)	(2.51)	(2.82)	(0.14)	(1.26)
Other Race	-0.547	**-3.533***	0.779	-0.162	-0.287
	(1.50)	(1.56)	(1.76)	(0.09)	(0.80)
Bilingual	0.004	0.335	0.644	0.009	-0.747
	(1.68)	(1.76)	(1.96)	(0.10)	(0.91)
Agency Selected	**4.934***	1.063	1.430	-0.038	2.104
	(2.49)	(2.65)	(2.94)	(0.15)	(1.37)
FPL <100%	0.741	4.428	-3.560	-0.057	**-2.636***
	(2.15)	(2.27)	(2.52)	(0.13)	(1.17)
FPL 100-300%	0.133	1.972	-1.138	-0.070	-0.286
	(1.67)	(1.78)	(1.99)	(0.10)	(0.91)
Income ≤20K	-2.686	-1.593	-1.563	0.038	-1.470
	(2.37)	(2.51)	(2.83)	(0.14)	(1.29)
Income 21K-40K	-1.419	-0.409	-0.862	0.006	-1.461
	(2.09)	(2.22)	(2.47)	(0.13)	(1.14)
Income 41K-60K	2.040	-1.865	-1.190	0.083	-0.500
	(2.13)	(2.27)	(2.55)	(0.13)	(1.17)
Income 61K-80K	**-3.578***	2.159	-1.625	0.039	0.160
	(1.77)	(1.88)	(2.09)	(0.11)	(0.96)
Class Size	0.360	-0.234	0.355	0.008	0.189
	(0.27)	(0.29)	(0.32)	(0.02)	(0.15)
CLASS_ES ≥5.5	-2.641	-2.394	1.807	-0.066	-0.153
	(2.38)	(2.47)	(2.79)	(0.14)	(1.27)
CLASS_CO ≥5.5	0.017	**3.012***	2.107	-0.039	-0.338
	(1.38)	(1.52)	(1.67)	(0.09)	(0.77)
CLASS_IS ≥3	1.432	-1.898	-0.676	-0.056	-0.051
	(1.20)	(1.29)	(1.43)	(0.07)	(0.66)
N	288	286	286	286	285

* p<0.05; ** p<0.01; *** p<0.001. Note: Reference groups omitted from the estimation are Males, White, English, FPL 300%+ and Income>80 thousand. Other controls are pre-test, age in months, days between tests, agencies, and an indicator for missing language or income. Raw scores are used for PPVT, and WJ or WM. Errors are clustered at the site level.

Appendix C.2. Sensitivity Analyses with DEEL Demographics.

Table C.2.1. Multivariate analyses of 2016–17 SPP program impacts, standard scores, including DEEL child characteristics

Table 20 Model	Rec. Vocabulary Standard	Literacy Standard	Math Standard	Executive Function	
				DCCS	PT
SPP	1.561	2.105	1.046	-0.130	**-1.314****
	(1.08)	(1.09)	(2.28)	(0.09)	(0.45)
Days Between Tests	0.048	0.044	0.148	-0.001	**0.088***
	(0.08)	(0.07)	(0.11)	(0.01)	(0.04)
Female	0.737	0.252	-1.995	0.095	-0.098
	(0.72)	(0.90)	(1.23)	(0.06)	(0.62)
Black	**-6.268*****	0.043	-2.832	-0.108	-0.009
	(1.55)	(1.87)	(2.35)	(0.09)	(0.81)
Asian	-3.502	-0.291	-1.247	-0.125	1.413
	(2.22)	(2.25)	(2.72)	(0.08)	(0.86)
Hispanic	**-5.974****	-3.757	-2.538	-0.217	-0.390
	(1.89)	(2.03)	(2.39)	(0.11)	(0.86)
Other	-2.340	-1.957	1.189	**-0.177***	0.056
	(1.27)	(1.40)	(1.46)	(0.08)	(0.65)
DLL	-0.689	2.751	0.664	-0.059	-0.855
	(1.90)	(2.28)	(2.13)	(0.07)	(0.76)
Agency Selected	1.170	0.122	0.785	-0.129	0.630
	(1.45)	(1.42)	(1.10)	(0.10)	(0.61)
Income ≤20K	**-3.954***	-1.204	-3.339	-0.060	**-3.406*****
	(1.90)	(2.08)	(2.24)	(0.14)	(0.57)
Income 21K-40K	-1.760	-1.173	-1.924	0.045	**-1.641***
	(1.73)	(1.89)	(1.91)	(0.11)	(0.68)
Income 41K-60K	1.406	-1.534	-2.257	-0.075	-1.404
	(1.63)	(1.53)	(2.21)	(0.13)	(0.88)
Income 61K-80K	**-5.466***	0.797	-2.744	-0.042	-0.698
	(2.27)	(1.74)	(1.37)	(0.10)	(0.84)
Observations	347	346	346	345	344

* p<0.05; ** p<0.01; *** p<0.001. Note: Reference groups omitted from the estimation are Males, White, English, FPL 300%+ and Income>80 thousand. Other controls are pre-test, age in months, days between tests, agencies, and an indicator for missing language or income. Errors are clustered at the site level.

Table C.2.2. Multivariate analyses of 2016–17 SPP program impacts, standard scores, including DEEL child characteristics, Model 1

Table 21 Model 1	Rec. Vocabulary Standard	Literacy Standard	Math Standard	Executive Function	
				DCCS	PT
SPP	0.750	1.696	-0.314	-0.091	**-1.295****
	(1.18)	(1.05)	(2.44)	(0.06)	(0.40)
Female	-0.240	0.201	**-2.105***	0.094	-0.102
	(0.80)	(0.82)	(0.96)	(0.05)	(0.50)
Black	**-6.663*****	0.753	-3.080	-0.103	-0.263
	(1.48)	(1.75)	(2.26)	(0.08)	(0.81)
Asian	**-3.932***	0.469	-1.376	-0.094	1.069
	(1.89)	(1.96)	(2.51)	(0.07)	(0.71)
Hispanic	**-6.052****	-2.309	-1.874	**-0.241***	-0.634
	(1.85)	(2.03)	(2.09)	(0.09)	(0.72)
Other	**-2.547***	-1.584	0.646	**-0.170***	-0.174
	(1.17)	(1.19)	(1.20)	(0.07)	(0.57)
DLL	-0.975	2.596	0.101	-0.064	-0.811
	(1.78)	(2.09)	(1.99)	(0.07)	(0.67)
Agency Selected	1.270	-0.114	0.851	-0.127	0.685
	(1.44)	(1.42)	(1.21)	(0.10)	(0.58)
Income ≤20K	-4.446*	-1.794	-3.656	-0.090	**-3.100*****
	(1.77)	(1.95)	(2.31)	(0.13)	(0.60)
Income 21K-40K	-2.350	-1.716	-2.161	0.001	-1.371
	(1.49)	(1.69)	(1.81)	(0.09)	(0.71)
Income 41K-60K	0.563	-2.177	-2.463	-0.114	-1.156
	(1.50)	(1.37)	(2.13)	(0.12)	(0.84)
Income 61K-80K	**-5.309***	-0.258	**-2.949***	-0.062	-0.415
	(2.17)	(1.49)	(1.27)	(0.08)	(0.79)
Observations	347	346	346	345	344

* p<0.05; ** p<0.01; *** p<0.001. Note: Reference groups omitted from the estimation are Males, White, English, FPL 300%+ and Income>80 thousand. Other controls are pre-test, age in months, days between tests, agencies, and an indicator for missing language or income. Errors are clustered at the site level.

Table C.2.3. Multivariate analyses of 2016–17 SPP program impacts, raw scores, including DEEL child characteristics, Model 2

Table 21 Model 2	Rec. Vocabulary Raw	Literacy Raw	Math Raw
SPP	1.729	0.725	0.199
	(1.40)	(0.37)	(0.76)
Female	1.142	0.337	-0.841
	(0.99)	(0.31)	(0.41)
Black	**-8.352****	0.103	-0.803
	(2.22)	(0.81)	(0.79)
Asian	-4.777	0.503	0.166
	(2.97)	(0.80)	(0.87)
Hispanic	**-8.419****	-1.181	-0.150
	(2.62)	(0.77)	(0.82)
Other	-2.751	-0.336	0.705
	(1.73)	(0.60)	(0.47)
DLL	-1.160	0.797	0.231
	(2.58)	(0.69)	(0.66)
Agency Selected	1.823	-0.149	0.439
	(1.98)	(0.59)	(0.33)
Income ≤20K	**-5.457***	-0.492	-1.334
	(2.47)	(0.80)	(0.70)
Income 21K-40K	-2.562	-0.138	-0.690
	(2.31)	(0.73)	(0.56)
Income 41K-60K	1.813	-0.409	-0.607
	(2.25)	(0.47)	(0.69)
Income 61K-80K	**-7.229***	0.580	**-0.919***
	(3.03)	(0.57)	(0.39)
Missing Income	-2.201	1.297	-1.959
	(4.48)	(0.96)	(1.14)
Observations	347	346	346

* p<0.05; ** p<0.01; *** p<0.001. Note: Reference groups omitted from the estimation are Males, White, English, FPL 300%+ and Income>80 thousand. Other controls are pre-test, age in months, days between tests, agencies, and an indicator for missing language or income. Errors are clustered at the site level.

Table C.2.4. Multivariate analyses of 2016–17 SPP program impacts, raw scores, including DEEL child characteristics, Model 3

Table 21 Model 3	Rec. Vocabulary Raw	Literacy Raw	Math Raw
SPP	1.142	**0.704***	-0.026
	(1.40)	(0.32)	(0.72)
Female	-0.095	0.372	**-0.779***
	(1.02)	(0.30)	(0.33)
Black	**-8.964*****	0.252	-1.054
	(2.12)	(0.72)	(0.74)
Asian	**-5.254***	0.614	-0.045
	(2.52)	(0.67)	(0.79)
Hispanic	**-8.536****	-0.750	-0.157
	(2.52)	(0.76)	(0.70)
Other	**-3.237***	-0.301	0.428
	(1.57)	(0.50)	(0.39)
DLL	-1.699	0.857	0.092
	(2.43)	(0.65)	(0.61)
Agency Selected	1.826	-0.200	0.465
	(1.99)	(0.58)	(0.40)
Income ≤20K	**-5.956***	-0.685	-1.388
	(2.32)	(0.76)	(0.75)
Income 21K-40K	-3.071	-0.305	-0.625
	(2.00)	(0.63)	(0.54)
Income 41K-60K	0.954	-0.634	-0.632
	(2.04)	(0.43)	(0.66)
Income 61K-80K	**-6.763***	0.254	**-0.885***
	(2.86)	(0.49)	(0.37)
Observations	347	346	346

* p<0.05; ** p<0.01; *** p<0.001. Note: Reference groups omitted from the estimation are Males, White, English, FPL 300%+ and Income>80 thousand. Other controls are pre-test, age in months, days between tests, agencies, and an indicator for missing language or income. Errors are clustered at the site level.

Table C.2.5. Multivariate analyses of 2016–17 SPP program impacts, standard scores, including DEEL child characteristics, Model 4

Table 21 Model 4	Rec. Vocabulary Standard	Literacy Standard	Math Standard	Executive Function	
				DCCS	PT
SPP	1.359	1.694	0.039	-0.123	**-1.330****
	(1.00)	(1.17)	(2.56)	(0.08)	(0.38)
Female	0.114	0.228	-2.165	0.085	0.033
	(0.80)	(0.88)	(1.14)	(0.06)	(0.59)
Black	**-6.184*****	0.526	-2.126	-0.102	-0.083
	(1.52)	(1.80)	(2.30)	(0.09)	(0.80)
Asian	-3.493	0.383	-0.255	-0.097	1.221
	(2.06)	(2.05)	(2.53)	(0.08)	(0.79)
Hispanic	**-5.581****	-2.600	-0.836	**-0.250***	-0.496
	(1.88)	(2.03)	(2.25)	(0.10)	(0.76)
Other	-2.212	-1.740	1.604	**-0.169***	0.020
	(1.21)	(1.28)	(1.40)	(0.08)	(0.61)
DLL	-0.469	2.378	-0.032	-0.070	-0.744
	(1.86)	(2.17)	(2.11)	(0.07)	(0.73)
Agency Selected	1.290	0.079	0.726	-0.126	0.652
	(1.42)	(1.47)	(1.09)	(0.10)	(0.59)
Income ≤20K	**-4.060***	-1.645	-3.667	-0.083	**-3.305*****
	(1.86)	(2.02)	(2.30)	(0.14)	(0.56)
Income 21K-40K	-2.025	-1.511	-2.071	0.012	**-1.591***
	(1.64)	(1.84)	(1.84)	(0.10)	(0.66)
Income 41K-60K	0.912	-2.004	-1.963	-0.108	-1.288
	(1.61)	(1.42)	(2.13)	(0.13)	(0.81)
Income 61K-80K	**-5.352***	0.307	**-2.756***	-0.057	-0.727
	(2.20)	(1.64)	(1.33)	(0.10)	(0.83)
Observations	347	346	346	345	344

* p<0.05; ** p<0.01; *** p<0.001. Note: Reference groups omitted from the estimation are Males, White, English, FPL 300%+ and Income>80 thousand. Other controls are pre-test, age in months, days between tests, agencies, and an indicator for missing language or income. Errors are clustered at the site level.

Table C.2.6. Multivariate analyses of 2016–17 SPP program impacts, raw scores, including DEEL child characteristics, Model 5

Table 21 Model 5	Rec. Vocabulary Raw	Literacy Raw	Math Raw
SPP	1.718	0.691	0.002
	(1.24)	(0.35)	(0.78)
Female	0.329	0.340	**-0.873***
	(1.07)	(0.31)	(0.39)
Black	**-8.373*****	0.214	-0.691
	(2.16)	(0.77)	(0.76)
Asian	-4.769	0.639	0.350
	(2.77)	(0.72)	(0.80)
Hispanic	**-7.965****	-0.783	0.235
	(2.59)	(0.79)	(0.74)
Other	-2.874	-0.308	0.766
	(1.66)	(0.56)	(0.45)
DLL	-0.988	0.717	0.065
	(2.52)	(0.67)	(0.65)
Agency Selected	1.922	-0.150	0.422
	(1.96)	(0.60)	(0.34)
Income ≤20K	**-5.487***	-0.616	-1.378
	(2.43)	(0.79)	(0.72)
Income 21K-40K	-2.744	-0.198	-0.647
	(2.19)	(0.70)	(0.55)
Income 41K-60K	1.276	-0.551	-0.491
	(2.22)	(0.44)	(0.66)
Income 61K-80K	**-6.879***	0.447	**-0.872***
	(2.97)	(0.54)	(0.37)
Observations	347	346	346

* p<0.05; ** p<0.01; *** p<0.001. Note: Reference groups omitted from the estimation are Males, White, English, FPL 300%+ and Income>80 thousand. Other controls are pre-test, age in months, days between tests, agencies, and an indicator for missing language or income. Errors are clustered at the site level.

Appendix D. Analyses for moderators.

Table D.1. Multivariate analyses of children's 2016–17 standard score gains in relation to child and classroom characteristics and parenting and school moderators: All variables and ECERS

Table 22	Rec. Vocabulary Standard	Literacy Standard	Math Standard	Executive Function	
				DCCS	PT
Female	1.100	0.622	-2.077	0.090	-0.616
	(0.98)	(1.04)	(1.15)	(0.06)	(0.53)
Black	**-4.765***	-0.565	-1.234	-0.032	-0.406
	(1.90)	(1.94)	(2.20)	(0.11)	(0.98)
Asian	-1.381	0.913	-0.097	-0.114	1.411
	(1.94)	(1.99)	(2.21)	(0.11)	(1.00)
Hispanic	**-5.897***	-4.505	-2.793	-0.171	-1.756
	(2.38)	(2.50)	(2.79)	(0.14)	(1.23)
Other	-0.568	**-3.217***	1.071	-0.143	-0.142
	(1.49)	(1.55)	(1.73)	(0.09)	(0.78)
DLL	0.188	-0.763	0.754	0.014	-0.471
	(1.74)	(1.81)	(2.00)	(0.10)	(0.92)
Agency Selected	**5.216***	0.627	2.166	-0.039	1.946
	(2.59)	(2.76)	(3.02)	(0.15)	(1.40)
Income ≤ 20K	-2.678	-1.831	-0.830	0.025	-1.075
	(2.39)	(2.52)	(2.81)	(0.14)	(1.28)
Income 21K-40K	-1.110	-0.878	-0.930	-0.018	-1.297
	(2.09)	(2.22)	(2.44)	(0.12)	(1.12)
Income 41K-60K	1.973	-2.410	-0.502	0.090	-0.395
	(2.16)	(2.30)	(2.54)	(0.13)	(1.16)
Income 61K-80K	**-3.563***	1.914	-1.182	0.015	0.192
	(1.78)	(1.88)	(2.07)	(0.11)	(0.95)
FPL <100%	0.429	4.304	-3.826	-0.032	**-2.729***
	(2.14)	(2.25)	(2.48)	(0.13)	(1.14)
FPL 100-300%	0.364	1.994	-0.996	-0.016	0.099
	(1.68)	(1.79)	(1.98)	(0.10)	(0.90)
Class Size	0.363	-0.195	0.412	0.007	0.172
	(0.27)	(0.29)	(0.32)	(0.02)	(0.15)
ECERS	0.438	1.871	2.455	-0.124	-0.793
	(1.44)	(1.52)	(1.66)	(0.08)	(0.76)
Positive Change	-0.120	0.020	-0.052	**-0.011***	**-0.124***
	(0.09)	(0.10)	(0.11)	(0.01)	(0.05)
Connection to preschool	0.104	0.207	0.090	**0.013***	0.074
	(0.11)	(0.11)	(0.13)	(0.01)	(0.06)
Teacher communication	0.408	-0.711	0.085	0.012	**0.479****
	(0.34)	(0.37)	(0.40)	(0.02)	(0.18)

Interaction with child	0.003	-0.074	-0.077	0.003	-0.029
	(0.05)	(0.05)	(0.05)	(0.00)	(0.03)
Previous center experience	0.447	-0.070	**2.896***	**0.161****	0.324
	(0.97)	(1.04)	(1.14)	(0.06)	(0.52)
Observations	288	286	286	286	285

* p<0.05; ** p<0.01; *** p<0.001. Note: Reference groups omitted from the estimation are Males, White, English, FPL 300%+ and Income>80 thousand. Other controls are pre-test, age in months, days between tests, agencies, and an indicator for missing language or income. Raw scores are used for PPVT, and WJ or WM.

Table D.2. Multivariate analyses of children's 2016–17 standard score gains in relation to child and classroom characteristics and parenting and school moderators: All variables and CLASS

Table 23	Rec. Vocabulary Standard	Literacy Standard	Math Standard	Executive Function	
				DCCS	PT
Female	0.985	0.554	-2.167	0.089	-0.624
	(0.97)	(1.04)	(1.16)	(0.06)	(0.53)
Black	**-4.347***	-0.640	-1.161	-0.022	-0.356
	(1.89)	(1.94)	(2.22)	(0.11)	(0.98)
Asian	-1.484	0.739	-0.126	-0.119	1.400
	(1.93)	(1.98)	(2.22)	(0.11)	(1.00)
Hispanic	**-5.673***	-4.535	-2.809	-0.156	-1.651
	(2.35)	(2.50)	(2.80)	(0.14)	(1.23)
Other	-0.497	**-3.351***	1.103	-0.149	-0.195
	(1.48)	(1.55)	(1.74)	(0.09)	(0.78)
DLL	0.014	-0.318	0.965	0.015	-0.436
	(1.71)	(1.80)	(2.00)	(0.10)	(0.91)
Agency Selected	**5.185***	0.594	1.917	-0.007	2.223
	(2.58)	(2.77)	(3.06)	(0.15)	(1.41)
Income ≤ 20K	-3.102	-1.365	-0.816	0.036	-0.989
	(2.37)	(2.52)	(2.83)	(0.14)	(1.27)
Income 21K-40K	-1.367	-0.649	-1.023	-0.006	-1.209
	(2.07)	(2.21)	(2.45)	(0.12)	(1.12)
Income 41K-60K	1.885	-2.162	-0.444	0.095	-0.396
	(2.14)	(2.30)	(2.56)	(0.13)	(1.16)
Income 61K-80K	**-3.649***	2.146	-1.161	0.022	0.247
	(1.76)	(1.87)	(2.07)	(0.11)	(0.95)
FPL <100%	1.166	4.180	-3.515	-0.045	**-2.845***
	(2.13)	(2.26)	(2.50)	(0.13)	(1.14)
FPL 100-300%	0.684	1.911	-0.678	-0.039	-0.039
	(1.67)	(1.79)	(1.99)	(0.10)	(0.90)
Class Size	0.316	-0.309	0.284	0.009	0.168
	(0.30)	(0.31)	(0.35)	(0.02)	(0.16)
CLASS_ES	-1.386	-1.335	0.179	**-0.152***	-0.961
	(1.27)	(1.34)	(1.49)	(0.07)	(0.68)
CLASS_CO	1.640	0.731	0.958	0.032	0.378
	(1.21)	(1.32)	(1.47)	(0.07)	(0.67)
CLASS_IS	1.007	-1.412	0.004	-0.052	-0.499
	(0.82)	(0.89)	(0.99)	(0.05)	(0.45)
Positive Change	-0.109	0.023	-0.053	**-0.011***	**-0.122***
	(0.09)	(0.10)	(0.11)	(0.01)	(0.05)
Connection to preschool	0.100	0.182	0.084	0.012	0.062
	(0.11)	(0.12)	(0.13)	(0.01)	(0.06)
Teacher communication	0.357	-0.704	0.108	0.008	**0.467***
	(0.34)	(0.37)	(0.40)	(0.02)	(0.18)

Interaction with child	0.001	-0.091	-0.085	0.003	-0.032
	(0.05)	(0.05)	(0.06)	(0.00)	(0.03)
Previous center experience	0.574	-0.143	**3.007****	**0.152****	0.253
	(0.96)	(1.04)	(1.15)	(0.06)	(0.52)
Observations	288	286	286	286	285

* p<0.05; ** p<0.01; *** p<0.001. Note: Reference groups omitted from the estimation are Males, White, English, FPL 300%+ and Income>80 thousand. Other controls are pre-test, age in months, days between tests, agencies, and an indicator for missing language or income. Raw scores are used for PPVT, and WJ or WM.

Table D.3. Multivariate analyses of 2016–17 SPP program impacts in relation to parenting and school moderators: Positive change

Table 24 Column 1 Model	Rec. Vocabulary Standard	Literacy Standard	Math Standard	Executive Function DCCS	Executive Function PT
SPP	1.414	0.399	-0.880	**-0.155***	**-1.169****
	(1.13)	(1.10)	(2.34)	(0.06)	(0.40)
Female	0.106	0.429	-2.020	0.090	0.006
	(0.79)	(0.87)	(1.17)	(0.06)	(0.59)
Age Month	0.022	**-0.230***	**-0.247***	**0.023*****	**0.198*****
	(0.09)	(0.08)	(0.11)	(0.00)	(0.05)
Black	**-6.195*****	0.766	-1.946	-0.095	-0.116
	(1.53)	(1.81)	(2.28)	(0.08)	(0.80)
Asian	-3.496	0.425	-0.226	-0.096	1.217
	(2.07)	(2.03)	(2.50)	(0.08)	(0.80)
Hispanic	**-5.594****	-2.362	-0.639	**-0.242***	-0.531
	(1.91)	(2.07)	(2.28)	(0.10)	(0.76)
Other	-2.217	-1.663	1.668	**-0.166***	0.010
	(1.21)	(1.36)	(1.43)	(0.08)	(0.62)
DLL	-0.485	2.692	0.191	-0.062	-0.785
	(1.85)	(2.14)	(2.13)	(0.07)	(0.73)
Agency Selected	1.315	-0.491	0.309	-0.141	0.726
	(1.44)	(1.44)	(1.03)	(0.10)	(0.61)
Income ≤ 20K	**-4.069***	-1.467	-3.527	-0.077	**-3.333*****
	(1.85)	(1.96)	(2.29)	(0.14)	(0.57)
Income 21K-40K	-2.016	-1.728	-2.220	0.007	-1.566*
	(1.66)	(1.90)	(1.85)	(0.10)	(0.66)
Income 41K-60K	0.920	-2.153	-2.067	-0.111	-1.269
	(1.60)	(1.45)	(2.14)	(0.13)	(0.80)
Income 61K-80K	**-5.351***	0.291	**-2.761***	-0.057	-0.727
	(2.21)	(1.67)	(1.30)	(0.10)	(0.83)
Positive Change	-0.005	**0.123****	0.087	0.003	-0.015
	(0.04)	(0.03)	(0.05)	(0.00)	(0.02)
Observations	366	365	365	364	363

* p<0.05; ** p<0.01; *** p<0.001. Note: Reference groups omitted from the estimation are Males, White, English, FPL 300%+ and Income>80 thousand. Other controls are pre-test, age in months, days between tests, agencies, and an indicator for missing language or income. Raw scores are used for PPVT, and WJ or WM.

Table D.4. Multivariate analyses of 2016–17 SPP program impacts in relation to parenting and school moderators: Connection to preschool

Table 24 Column 2 Model	Rec. Vocabulary Standard	Literacy Standard	Math Standard	Executive Function	
				DCCS	PT
SPP	1.118	-0.136	-1.095	**-0.195****	**-1.380****
	(1.15)	(1.07)	(2.33)	(0.07)	(0.43)
Female	0.160	0.564	-1.959	0.100	0.042
	(0.79)	(0.87)	(1.17)	(0.06)	(0.59)
Age Month	0.027	**-0.221***	**-0.243***	**0.024*****	**0.202*****
	(0.09)	(0.08)	(0.12)	(0.00)	(0.05)
Black	**-6.160*****	0.673	-2.058	-0.093	-0.078
	(1.51)	(1.78)	(2.31)	(0.08)	(0.80)
Asian	-3.476	0.485	-0.212	-0.093	1.224
	(2.05)	(1.98)	(2.53)	(0.08)	(0.79)
Hispanic	**-5.519****	-2.203	-0.599	**-0.230***	-0.484
	(1.85)	(2.05)	(2.29)	(0.10)	(0.75)
Other	-2.190	-1.593	1.685	**-0.162***	0.024
	(1.19)	(1.32)	(1.42)	(0.08)	(0.61)
DLL	-0.467	2.366	-0.049	-0.070	-0.744
	(1.86)	(2.09)	(2.10)	(0.07)	(0.74)
Agency Selected	1.172	-0.808	0.169	-0.162	0.627
	(1.45)	(1.47)	(1.11)	(0.10)	(0.62)
Income ≤ 20K	**-4.031***	-1.470	-3.577	-0.073	**-3.299*****
	(1.85)	(1.96)	(2.28)	(0.14)	(0.57)
Income 21K-40K	-2.051	-1.709	-2.201	0.006	**-1.595***
	(1.64)	(1.91)	(1.89)	(0.10)	(0.66)
Income 41K-60K	0.877	-2.239	-2.127	-0.116	-1.294
	(1.62)	(1.43)	(2.19)	(0.13)	(0.81)
Income 61K-80K	**-5.378***	0.097	-2.887*	-0.065	-0.732
	(2.20)	(1.71)	(1.29)	(0.10)	(0.83)
Connection to preschool	0.023	**0.175*****	0.108	**0.007*****	0.005
	(0.03)	(0.04)	(0.07)	(0.00)	(0.02)
Observations	366	365	365	364	363

* p<0.05; ** p<0.01; *** p<0.001. Note: Reference groups omitted from the estimation are Males, White, English, FPL 300%+ and Income>80 thousand. Other controls are pre-test, age in months, days between tests, agencies, and an indicator for missing language or income. Raw scores are used for PPVT, and WJ or WM.

Table D.5. Multivariate analyses of 2016–17 SPP program impacts in relation to parenting and school moderators: Teacher communication

Table 24 Column 3 Model	Rec. Vocabulary Standard	Literacy Standard	Math Standard	Executive Function	
				DCCS	PT
SPP	1.015	0.932	-0.363	-0.143	**-1.668*****
	(1.17)	(1.14)	(2.54)	(0.07)	(0.42)
Female	0.117	0.249	-2.153	0.085	0.038
	(0.80)	(0.89)	(1.15)	(0.06)	(0.58)
Age Month	0.027	**-0.244****	**-0.258***	**0.023*****	**0.205*****
	(0.09)	(0.08)	(0.12)	(0.00)	(0.05)
Black	**-6.102*****	0.687	-2.044	-0.098	-0.001
	(1.52)	(1.83)	(2.29)	(0.09)	(0.81)
Asian	-3.476	0.402	-0.252	-0.097	1.232
	(2.02)	(2.02)	(2.54)	(0.08)	(0.76)
Hispanic	**-5.538****	-2.569	-0.808	**-0.248***	-0.460
	(1.86)	(1.97)	(2.27)	(0.10)	(0.77)
Other	-2.135	-1.610	1.674	**-0.165***	0.093
	(1.20)	(1.35)	(1.43)	(0.08)	(0.61)
DLL	-0.276	2.806	0.190	-0.059	-0.560
	(1.89)	(2.15)	(2.13)	(0.07)	(0.75)
Agency Selected	1.136	-0.266	0.535	-0.136	0.495
	(1.47)	(1.36)	(1.14)	(0.10)	(0.63)
Income ≤ 20K	**-3.959***	-1.458	-3.567	-0.077	**-3.209*****
	(1.85)	(2.03)	(2.32)	(0.14)	(0.61)
Income 21K-40K	-2.023	-1.519	-2.075	0.012	-1.588*
	(1.65)	(1.92)	(1.87)	(0.10)	(0.66)
Income 41K-60K	0.894	-2.076	-2.006	-0.109	-1.299
	(1.64)	(1.48)	(2.16)	(0.13)	(0.84)
Income 61K-80K	**-5.379***	0.232	**-2.793***	-0.058	-0.753
	(2.22)	(1.67)	(1.30)	(0.10)	(0.84)
Teacher communication	0.166	0.378	0.199	0.010	**0.164***
	(0.16)	(0.19)	(0.32)	(0.01)	(0.08)
Observations	366	365	365	364	363

* p<0.05; ** p<0.01; *** p<0.001. Note: Reference groups omitted from the estimation are Males, White, English, FPL 300%+ and Income>80 thousand. Other controls are pre-test, age in months, days between tests, agencies, and an indicator for missing language or income. Raw scores are used for PPVT, and WJ or WM.

Table D.6. Multivariate analyses of 2016–17 SPP program impacts in relation to parenting and school moderators: Interactions with child

Table 24 Column 4 Model	Rec. Vocabulary Standard	Literacy Standard	Math Standard	Executive Function	
				DCCS	PT
SPP	1.345	1.621	-0.215	-0.119	**-1.434*****
	(0.97)	(1.18)	(2.60)	(0.08)	(0.38)
Female	0.112	0.210	-2.230	0.086	0.013
	(0.80)	(0.88)	(1.13)	(0.06)	(0.57)
Age Month	0.023	**-0.255****	**-0.267***	**0.023*****	**0.203*****
	(0.08)	(0.08)	(0.11)	(0.00)	(0.05)
Black	**-6.168*****	0.592	-1.934	-0.105	-0.004
	(1.56)	(1.78)	(2.29)	(0.09)	(0.80)
Asian	-3.484	0.420	-0.145	-0.099	1.280
	(2.07)	(2.04)	(2.50)	(0.08)	(0.78)
Hispanic	**-5.603****	-2.711	-1.248	**-0.242***	-0.666
	(1.87)	(2.10)	(2.32)	(0.10)	(0.73)
Other	-2.201	-1.681	1.784	**-0.171***	0.095
	(1.23)	(1.28)	(1.39)	(0.08)	(0.61)
DLL	-0.478	2.331	-0.197	-0.068	-0.818
	(1.87)	(2.17)	(2.15)	(0.07)	(0.73)
Agency Selected	1.293	0.093	0.773	-0.127	0.676
	(1.42)	(1.46)	(1.12)	(0.10)	(0.62)
Income ≤ 20K	**-4.039***	-1.554	-3.388	-0.087	**-3.209*****
	(1.89)	(1.99)	(2.31)	(0.14)	(0.55)
Income 21K-40K	-2.011	-1.448	-1.869	0.009	**-1.519***
	(1.66)	(1.82)	(1.81)	(0.10)	(0.66)
Income 41K-60K	0.924	-1.934	-1.742	-0.110	-1.223
	(1.63)	(1.41)	(2.12)	(0.13)	(0.81)
Income 61K-80K	**-5.327***	0.428	-2.355	-0.064	-0.567
	(2.21)	(1.64)	(1.42)	(0.10)	(0.85)
Interactions with child	-0.006	-0.027	**-0.092***	0.002	-0.040
	(0.03)	(0.04)	(0.04)	(0.00)	(0.02)
Observations	366	365	365	364	363

* p<0.05; ** p<0.01; *** p<0.001. Note: Reference groups omitted from the estimation are Males, White, English, FPL 300%+ and Income>80 thousand. Other controls are pre-test, age in months, days between tests, agencies, and an indicator for missing language or income. Raw scores are used for PPVT, and WJ or WM.

Table D.7. Multivariate analyses of 2016–17 SPP program impacts in relation to parenting and school moderators: Previous center experience

Table 24 Column 5 Model	Rec. Vocabulary Standard	Literacy Standard	Math Standard	Executive Function	
				DCCS	PT
SPP	1.469	1.652	-0.049	-0.130	**-1.325****
	(1.00)	(1.17)	(2.71)	(0.08)	(0.40)
Female	0.033	0.262	-2.088	0.090	0.029
	(0.81)	(0.87)	(1.12)	(0.06)	(0.59)
Age Month	0.018	**-0.252****	**-0.260***	**0.023*****	**0.201*****
	(0.09)	(0.08)	(0.12)	(0.00)	(0.05)
Black	**-6.229*****	0.529	-2.105	-0.100	-0.084
	(1.54)	(1.80)	(2.30)	(0.09)	(0.80)
Asian	-3.663	0.448	-0.105	-0.087	1.214
	(2.07)	(2.06)	(2.52)	(0.08)	(0.80)
Hispanic	**-5.640****	-2.585	-0.788	**-0.246***	-0.498
	(1.85)	(2.05)	(2.24)	(0.10)	(0.75)
Other	-2.271	-1.723	1.650	**-0.165***	0.018
	(1.19)	(1.29)	(1.34)	(0.08)	(0.62)
Missing Race	-0.276	4.921	4.372	-0.161	**-4.508***
	(3.66)	(2.55)	(3.39)	(0.14)	(1.88)
DLL	-0.403	2.360	-0.068	-0.073	-0.742
	(1.86)	(2.14)	(2.09)	(0.07)	(0.74)
Agency Selected	1.352	0.059	0.681	-0.129	0.655
	(1.31)	(1.48)	(1.16)	(0.10)	(0.59)
Income ≤ 20K	**-4.102***	-1.617	-3.590	-0.077	**-3.309*****
	(1.88)	(2.04)	(2.30)	(0.14)	(0.56)
Income 21K-40K	-1.920	-1.552	-2.154	0.007	**-1.587***
	(1.62)	(1.85)	(1.85)	(0.10)	(0.66)
Income 41K-60K	0.662	-1.918	-1.764	-0.093	-1.298
	(1.68)	(1.43)	(2.15)	(0.13)	(0.82)
Income 61K-80K	**-5.487***	0.364	-2.627	-0.049	-0.734
	(2.28)	(1.64)	(1.36)	(0.09)	(0.84)
Previous center experience	-1.043	0.437	0.962	0.064	-0.047
	(0.86)	(0.83)	(1.29)	(0.05)	(0.43)
Observations	366	365	365	364	363

* p<0.05; ** p<0.01; *** p<0.001. Note: Reference groups omitted from the estimation are Males, White, English, FPL 300%+ and Income>80 thousand. Other controls are pre-test, age in months, days between tests, agencies, and an indicator for missing language or income. Raw scores are used for PPVT, and WJ or WM.

Table D.8. Multivariate analyses of 2016–17 SPP program impacts in relation to parenting and school moderators: All variables

Table 24 Column 6 Model	Rec. Vocabulary Standard	Literacy Standard	Math Standard	Executive Function	
				DCCS	PT
SPP	1.120	-0.233	-1.869	**-0.181***	**-1.717*****
	(1.08)	(1.13)	(2.63)	(0.07)	(0.41)
Female	0.030	0.564	-1.881	0.109	-0.043
	(0.81)	(0.88)	(1.11)	(0.06)	(0.55)
Age Month	0.022	**-0.217***	**-0.239***	**0.023*****	**0.203*****
	(0.09)	(0.08)	(0.12)	(0.00)	(0.05)
Black	**-6.145*****	0.877	-1.835	-0.108	0.050
	(1.61)	(1.77)	(2.31)	(0.08)	(0.82)
Asian	-3.609	0.660	0.024	-0.085	1.291
	(2.04)	(1.96)	(2.48)	(0.08)	(0.76)
Hispanic	**-5.624****	-2.178	-1.095	**-0.227***	-0.729
	(1.79)	(2.29)	(2.37)	(0.10)	(0.69)
Other	-2.147	-1.485	1.941	**-0.164***	0.229
	(1.22)	(1.32)	(1.37)	(0.08)	(0.62)
DLL	-0.331	2.104	-0.344	-0.103	-0.613
	(1.81)	(2.08)	(2.19)	(0.07)	(0.74)
Agency Selected	1.148	-1.056	0.181	-0.149	0.577
	(1.47)	(1.47)	(1.18)	(0.10)	(0.72)
Income ≤ 20K	**-3.968***	-1.373	-3.128	-0.076	**-3.076*****
	(1.89)	(2.00)	(2.33)	(0.14)	(0.63)
Income 21K-40K	-1.840	-1.641	-2.081	0.003	-1.418*
	(1.65)	(1.95)	(1.88)	(0.11)	(0.66)
Income 41K-60K	0.686	-1.977	-1.803	-0.106	-1.206
	(1.77)	(1.36)	(2.22)	(0.13)	(0.86)
Income 61K-80K	**-5.526***	0.357	-2.295	-0.067	-0.624
	(2.31)	(1.78)	(1.42)	(0.09)	(0.88)
Positive Change	-0.060	0.015	0.021	-0.006	-0.064
	(0.06)	(0.06)	(0.09)	(0.00)	(0.04)
Connection to preschool	0.057	**0.260****	0.101	**0.012***	0.015
	(0.10)	(0.09)	(0.11)	(0.01)	(0.06)
Teacher communication	0.228	-0.113	-0.131	-0.016	**0.331***
	(0.20)	(0.24)	(0.38)	(0.02)	(0.14)
Interaction with child	-0.012	-0.056	**-0.115***	0.001	-0.045
	(0.03)	(0.03)	(0.04)	(0.00)	(0.02)
Previous center experience	-1.033	0.245	0.689	0.059	-0.048
	(0.90)	(0.83)	(1.26)	(0.05)	(0.44)
Observations	366	365	365	364	363

* p<0.05; ** p<0.01; *** p<0.001. Note: Reference groups omitted from the estimation are Males, White, English, FPL 300%+ and Income>80 thousand. Other controls are pre-test, age in months, days between tests, agencies, and an indicator for missing language or income. Raw scores are used for PPVT, and WJ or WM.

Appendix E. Indicators tables for interactions with children.

ECERS-3 Interaction indicators		2016		2017	
		N	Percent	N	Percent
Item 5. Child Related Display	Indicator 5.4 Staff talk about display materials at least two different times during free play and/or routines in a way that interests the children	2	14.3	3	9.38
	Indicator 7.4 Staff are observed pointing out and reading the words in the display in a way that interests the children.	1	7.1	17	53.13
Item 13. Encouraging children to use language	Indicator 1.4 Staff make no attempt to encourage children to communicate (Ex: no singing, nursery rhymes, saying alphabet, naming colors).	14	100	32	100
	Indicator 1.5 Social environment does not encourage much talking among children or with staff (Ex: strict atmosphere where child talk not encouraged; little time to interact socially.	14	100	32	100
	Indicator 5.3 Staff respond positively to children's communication and encourage them to talk more.	11	78.6	30	93.75
	Indicator 5.4- Staff help children communicate verbally with one another.	12	85.7	25	78.13
	Indicator 7.3. Staff-child conversations go beyond classroom activities and materials (Ex: include social talk about home and family life; activities in the community; feelings; other non-school topics).	6	42.9	21	65.63
Item 14. Staff use of books with children	Indicator 1.2. Book times are unpleasant or not engaging for many of the children (Ex: children forced to listen; punitive atmosphere; children can't see book; children's reactions are treated as interruptions).	9	64.3	25	78.13
	Indicator 1.3. Staff reading or use of books with children is dull, disinterested, and/or unenthusiastic.	9	64.3	25	78.13
	Indicator 3.2- Book time is arranged to encourage children's engagement (Ex: children can easily see the book; crowding does not cause problems; books used that interest children; appropriate length).	9	64.3	25	78.13
	Indicator 3.3- The majority of children appear to be engaged for most of the time when books are used (Ex: children may lose interest for short period, but then become interested again; one child is not interested but others are).	9	64.3	25	78.13
	Indicator 5.3- All children participating in the activity are actively engaged during each book time (Ex: staff is supportive and reads with interest; children appear to enjoy book time and pay attention).	5	35.7	10	31.25
	Indicator 7.2- Staff and children discuss the content of a book in a way that engages children.	1	7.1	18	56.25
Item 15. Encouraging children's use of books	Indicator 5.2- Children show interest in accessible books (Ex: child chooses to use books in the cozy area during free play; looks at book in science center).	9	64.3	25	78.13

	Indicator 5.3- Books are organized in a defined reading interest center, with a place to store the books for easy access and a space with comfortable furnishing to use them.	9	64.3	30	93.75
	Indicator 7.3- Most accessible books are displayed in order to encourage book use (Ex: books not crowded on shelf; many covers easily seen).	5	35.7	24	75
Item 16. Print	Indicator 1.2- Staff respond negatively when children show little or no interest in activities used to teach letters or words (Ex: scold child or send to time-out, make child work on letter activity until finished even though others get to play).	14	100	32	100
Item 17. Fine motor	Indicator 5.3- Staff show some interest as children use fine motor materials (Ex: ask short answer questions about color or shape; participate in play).	14	100	31	96.88
	Indicator 7.1- Staff show more extended interest in what children create/do with the materials (Ex: have conversations with children about what they make; show how to use materials; have children select materials of appropriate interest & difficulty).	5	35.7	20	62.5
Item 18. Art	Indicator 3.2- Some individual expression with art materials is observed as children use art materials, or observed in the display (Ex: Allowed to do free drawing; paints at easel for child to use in own way; play dough used without cookie cutters).	14	100	32	100
	Indicator 3.3- Some positive staff involvement with children using art materials (Ex: staff make comments to show appreciation about a child's work; identify colors or shapes seen in a child's creation).	14	100	31	96.88
	Indicator 5.3- Staff have conversations with interested children about their work (Ex: "Tell me about your picture." "How did you make that clay form?").	7	50	19	59.38
	Indicator 7.3- Staff write captions dictated by interested children about their artwork or help older children to write captions for themselves if they wish to (Ex: "You said, 'This is my new puppy.' See, I wrote your words.").	3	21.4	7	21.88
Item 19. Music & mov.	Indicator 3.4- Staff-led group music activities are pleasant and children generally appear to be engaged.	13	92.9	24	75
Item 20. Blocks	Indicator 5.5- Staff have many conversations with interested children about their block play (Ex: ask questions about what children are building or their favorite shapes to use; talk about pictures of structures with the children).	3	21.4	9	28.13
Item 21. Dramatic play	Indicator 1.3- Staff usually ignore children in the dramatic play area, except to stop disruptive behavior (Ex: staff settle conflicts, manage rotation of turns, or ask children to lower their voices).	10	71.4	27	84.38
	Indicator 5.3- Staff carry on conversations with the children as they play, joining in but not taking over (Ex: relate children's play to their home experiences; discuss the roles children are playing; encourage play based on field trip).	5	35.7	15	46.88

	Indicator 7.2- Staff talk with children about print and numbers in dramatic play in a way that is meaningful to the children (Ex: discuss menus with prices for restaurants; help children make signs and price tags for store play).	1	7.1	3	9.38
Item 22. Nature/science	Indicator 1.2- Staff do not talk about nature/science with the children during the observation (Ex: mention weather, seasons; read factual book on animals; mention temperature of water).	11	78.6	26	81.25
Item 23. Math mats. & act.	Indicator 1.3- Most math activities do not keep most children engaged (Ex: not appropriate for developmental level; children frequently lose attention or are frustrated with activities; children rarely select math activities independently).	14	100	32	100
	Indicator 3.3- Math activities used engage most of the participating children (Ex: most children are interested in calendar activities that are math related; enjoy rote counting at group time).	11	78.6	21	65.63
Item 24. Math in daily events	Indicator 5.1- Staff encourage math learning as part of daily routines (Ex: explain setting table; name rectangular and round tables when saying where to put plates and cups; counting to 20 while washing hands).	3	21.4	10	31.25
	Indicator 5.2- Staff engage children in conversations about math as they play in non-math areas (Ex: discuss using measuring cups to water plant; count how many teacups are needed for dolls; talk about how to measure feet in play shoe store).	2	14.3	12	37.5
Item 25. Written numbers	Indicator 3.3- When children play with materials credited in 3.2 (ex: play money), staff sometimes point out the numbers and talk about them in a way that interests children.	3	21.4	8	25
Item 27. Technology	Indicator 5.4- Staff are actively involved with children in use of electronic media (Ex: do activity suggested in educational TV program; help child learn to use computer program).	2	14.3	1	3.13
Item 28. Gross motor	Indicator 3.3- Staff show some interest in children's gross motor activity (Ex: make sure children get scheduled gross motor times; encourage children to run or climb; respond when child calls for attention in gross motor activities).	11	78.6	29	90.63
	Indicator 5.3- Staff show much interest in children who participate in gross motor activity (Ex: do not pay most attention to children during sedentary activities; show enthusiasm when children run, slide, jump; help children learn to use equipment).	8	57.1	22	68.75
Item 29. Individualized T&L	Indicator 1.1- Almost all teaching uses a one-size-fits-all approach (Ex: all children must do the same activity in the same way; expectations are not based on children's individual	14	100	32	100
	Indicator 3.1- Some teaching uses an individualized approach (Ex: responds to individual interests during circle or meal times).	14	100	31	96.88
	Indicator 5.2- Staff sometimes circulate through classroom, adding individualized learning to children's activities (Ex: counts blocks with child who built a tower; shows child how to play sorting games).	5	35.7	22	68.75

	Indicator 7.1- Most teaching is individualized, with few if any exceptions.	1	7.14	12	37.5
	Indicator 7.2- Much individualized teaching while children participate in free play (Ex: staff circulate often to various areas of room; children's play is enhanced and not interrupted when teaching occurs).	2	14.3	11	34.38
Item 31. Peer interaction	Indicator 3.1- Children have some time to select their own companions and activities during the observation (Ex: some free play is observed, indoors or outdoors).	14	100	32	100
	Indicator 5.2- Staff generally help the children solve social problems in a satisfying way (Ex: help children take turns with a tricycle; help shy child find a chair to join in an art activity).	13	92.9	28	87.5
Item 33. Transitions	Indicator 1.2- Staff usually not prepared for what comes next in the schedule.	14	100	32	100
	Indicator 1.4- Children required to wait for 10 minutes or more during any transition, with nothing engaging to do (Ex: waiting at table to eat; waiting in line; waiting for teacher to begin circle time).	14	100	32	100
	Indicator 7.1- Transitions are often gradual or individualized (Ex: children can go outside while others are still getting ready; children can begin eating as soon as they sit at table; circle time begins while some children are still cleaning up).	5	35.7	19	59.38
Item 34. Free play	Indicator 7.2- Staff use a wide variety of words to expand children's knowledge during free play activities.	1	7.14	8	25
Item 35. Whole-group activities	Indicator 5.3- Staff use group times to introduce children to meaningful ideas in which children are interested (Ex: review theme of the week; explain how to use new material; tell children what will happen on field trip).	8	57.1	25	78.13
	Indicator 7.1- All children in the group are actively engaged in group activities.	4	28.6	4	12.5

Appendix F. Family Survey Tables.

This appendix first compares survey non-respondents and respondents based on DEEL data on gender, ethnicity, language and FPL for all children on the target sample of 320. It then compares survey non-respondents and respondents for the 291 children for which we were able to collect pre- and post-test data on at least one measure. Groups of respondents and non-respondents differed in ethnicity and FPL. The rest of the appendix presents tables for respondents the different indicators and information captured by the family survey for respondents enrolled in SPP.

Table F.1.a, Respondents (45.5% of children's families), non-Respondents and Non-Consented from target sample

DEEL Child Information		Respondent %	Non-respondent %
Total (N=320)			
Gender	Male (N=157)	50.8	51.60
	Female (N=163)	49.2	48.40
Ethnicity*	White (N=69)	24.4	9.70
	Black (N=72)	19.4	35.50
	Asian (N=53)	17.4	12.9
	Hispanic (N=44)	12.4	19.4
	Other (N=79)	26.4	17.7
	Unknown (N=3)	0.0	4.8
Language*	English (N=207)	80.2	0.0
	Spanish (N=3)	1.2	0.0
	Vietnamese (N=12)	4.7	0.0
	Other (N=35)	13.6	0.0
	Unknown (N=63)	0.4	100.0
FPL*	<100 (N=70)	19.0	33.9
	100-300 (N=150)	45.3	53.2
	>300 (N=98)	35.7	9.7
	Unknown (N=2)	0.0	3.2

*Respondent versus Non-respondent distribution was statistically significantly different.

Table F.1.b. Respondents (53.1% of children's families) versus Non-respondents, children with pre- and post-test

DEEL Child Information		Respondent %	Non-respondent %
Total (N=291)			
Gender	Male (N=150)	51.6	51.1
	Female (N=141)	48.4	48.9
Ethnicity*	White (N=65)	25.2	6.7
	Black (N=63)	18.3	40.0
	Asian (N=48)	17.1	13.3
	Hispanic (N=43)	13.0	24.4
	Other (N=72)	26.4	15.6
	Unknown (N=0)	0.0	0.0
Language*	English (N=199)	80.9	0.0
	Spanish (N=3)	1.2	0.0
	Vietnamese (N=12)	4.9	0.0
	Other (N=31)	12.6	0.0
	Unknown (N=46)	0.4	100.0
FPL*	<100 (N=57)	18.7	33.3
	100-300 (N=80)	45.1	57.8
	>300 (N=41)	36.2	8.9
	Unknown (N=0)	0.0	0.0

*Respondent versus Non-respondent distribution was statistically significantly different.

Socioeconomic indicators

Table F.2. Socioeconomic indicators

Parent Education	N	Percent
High school diploma	31	12.25%
Some college	63	24.90%
Associate's degree	32	12.65%
Bachelor's degree	81	32.02%
Master's degree or higher	46	18.18%
Total	253	100.00%
Annual household income	N	Percent
20,000 or less	44	17.40%
21,000-40,000	56	22.13%
41,000-60,000	46	18.19%
61,000-80,000	45	17.78%
81,000	62	24.51%
Total	253	100%
Family Structure	N	Percent
Two parents (both biological or adoptive)	180	71.43%
Two parents (one biological and one other)	9	3.57%
One parent	60	23.81%
Other	3	1.19%
Total	252	100.00%

OTHERS	Mean	SD
Years in current residence	3.279	1.844
Age of mother at birth/adoption of child	30.73469	6.717287

Welfare

Table F.3. Welfare

Public Benefit	N	Percent	Total
Food stamps	60	23.90%	251
WIC	53	21.12%	251
TANF	8	3.24%	247
Early Head Start	4	1.63%	246
Head Start	18	7.26%	248
Medicaid	68	27.64%	246
Medicare	38	15.51%	245
ESEAP	13	5.31%	245
Working Connections	19	7.69%	247
Food Bank	15	6.07%	247

Language and Immigration

Table F.4. Home language

Primary Language	N	Percent
English	204	80.31%
Spanish	3	1.18%
Vietnamese	12	4.72%
Chinese	9	3.54%
Other	26	10.22%
Total	254	100.00%

Preschool Choices

Table F.5. Preschool choices

Importance if cost was not an issue	N	Percent	Total
Focus on social and emotional development	118	45.91%	257
Location	89	34.63%	257
What is taught and how is taught	89	34.63%	257
Focus on kindergarten readiness	87	33.85%	257
Focus on academic skills	69	26.85%	257
Diversity of students	61	23.74%	257
Hours of operation	48	18.68%	257
Dual-language program	36	14.01%	257
Adult:child ratio	34	13.23%	257
Teacher experience	33	12.84%	257
Focus on outdoor play and nature	30	11.67%	257

Reputation	25	9.73%	257
Diversity of staff	18	7.00%	257
Teacher education	16	6.23%	257
Located in neighborhood elementary school	15	5.84%	257
Early Achiever rating	10	3.89%	257
Special programming	8	3.11%	257

Parental perceptions on SPP programs and teachers

Table F.6. Perceptions on Positive changes on child since SPP enrollment

	N	Strongly disagree	2	3	4	5	Strongly agree
Language	257	0.78%	1.17%	13.23%	17.90%	29.96%	36.96%
Physical Development	256	1.17%	0.39%	16.41%	21.88%	35.55%	24.61%
Behavioral/ Socio-Emotional	253	1.19%	4.74%	12.65%	21.34%	32.81%	27.27%
Literacy	255	1.57%	5.49%	18.04%	20.78%	27.84%	26.27%
Math	251	5.58%	8.37%	21.12%	20.32%	24.70%	19.92%
Science	250	6.80%	9.60%	22%	23.60%	21.20%	16.80%

Table F.7. Perceptions of the teacher

Regarding the child's teacher	N	No	Yes
Talks to me each day	255	12.55%	83.92%
Uses a curriculum for teaching	254	24.80%	75.20%
Teaches my child behavioral/social/emotional skills	253	1.98%	90.12%
Teaches my child academic skills	256	1.17%	85.16%
Tracks my child's progress	254	1.97%	82.28%
Is fluent in my child's primary home language	256	12.11%	81.25%
Has a Bachelor's degree	255	1.57%	50.98%
Engages in training opportunities	255	1.18%	61.96%

Table F.8. Perceptions of feeling welcome or unwelcome by the program

How much do you agree: I feel welcome at the preschool	N	Percent
Strongly disagree	1	0.39%
2	3	1.17%
3	6	2.34%
4	16	6.25%
5	45	17.58%
Strongly agree	185	72.27%
Total	256	100%

Table F.9. Regarding the program

Regarding the child's program	N	Strongly disagree	2	3	4	5	Strongly agree
I feel connected with my child's teacher	256	1.17%	3.13%	8.59%	16.80%	28.13%	42.19%
I feel connected with my child's preschool	255	1.96%	3.92%	7.84%	20%	28.63%	37.65%
I have received work samples	253	1.19%	1.19%	6.32%	6.72%	23.72%	60.87%
I have received assessment results	256	2.34%	1.56%	7.42%	14.84%	27.34%	46.48%
I know about the curriculum that is used	256	3.13%	5.47%	13.67%	22.66%	25.39%	29.69%
I feel welcome at the preschool	256	0.39%	1.17%	2.34%	6.25%	17.58%	72.27%
I have received feedback about my child's performance	256	1.17%	1.95%	5.86%	10.94%	25.78%	54.30%

Parenting practices

Table F.10. Parenting activities with the child

In a typical week, how often do you…	Total	Not at all	1-2 times per week	3-6 times per week	Every day
Play toys with your child	251	2.39%	32.67%	33.07%	31.87%
Tell stories to your child	254	5.12%	28.35%	29.53%	37.01%
Sing songs and/or dance with your child	254	3.54%	19.29%	37.80%	39.37%
Help your child to do arts & crafts	254	6.30%	38.19%	37.80%	17.72%
Write with your child	254	9.06%	41.73%	33.86%	15.35%
Involve your child in household chores	254	3.54%	17.32%	36.22%	42.91%
Take your child on errands	254	0.39%	24.41%	38.58%	36.61%
Play pretend or role-playing games	254	4.72%	38.19%	33.46%	23.62%
Watch TV with your child	253	4.74%	35.18%	29.25%	30.83%
Play video games with your child	253	58.89%	26.88%	11.07%	3.16%
Do puzzles with your child	253	19.37%	54.94%	20.16%	5.53%
Talk about numbers and/or shapes with your child	253	1.58%	22.53%	42.69%	33.20%
Talk about nature or do science projects with your child	255	11.37%	41.57%	31.76%	15.29%
Build or play construction toys with your child	255	10.20%	44.71%	29.41%	15.69%
Take your child to the library	251	33.47%	58.17%	6.37%	1.99%

Go for a walk/play outside with your child	253	1.98%	46.25%	38.34%	13.44%
Take your child to the park or playground	251	1.99%	66.53%	26.29%	5.18%
Take your child to museum/zoo/other ed. site	250	29.60%	63.20%	5.20%	2%
Play a sport or exercise together	251	25.10%	51.39%	15.54%	7.97%
Engage in faith-based activities	253	49.41%	29.25%	11.86%	9.49%
Visit relatives or friends	254	7.87%	59.84%	24.02%	8.27%
Extra academic program	249	88.76%	8.43%	1.61%	1.20%
Play board or card games with your child	252	19.05%	52.78%	21.83%	6.35%

Table F.11. Number of books in the home

Number of book in the home	N	Percent
less than 20 books	39	15.42%
20-50 books	89	35.18%
More than 50 books	125	49.41%
Total	257	100.00%

Other care and past care

Table F.12. Out-of-home care used in addition to SPP

Attend any other out-of-home care in addition to SPP?	28	10.89%
Extended day child care	9	34.62%
Developmental preschool	2	7.69%
With a relative	8	30.77%
With a friend or neighbor	2	7.69%
Childcare someplace else	5	19.23%
Total	26	100%

Table F.13. Expulsions

Child asked to leave a childcare or preschool because of behavior?	N	Percent	Total
Yes	7	2.72%	257
One time	3	1.17%	5
Two times	1	0.39%	5
Three times	1	0.39%	5

Table F.14. Out-of-home care used prior to SPP

Experiences prior to SPP	Birth-1	1-2 yr. old	2-3 yr. old	3-4 yr. old
At home with parent or family member	75.00%	62.67%	44.30%	32.06%
At home under non-relative care	7.00%	8.67%	4.43%	4.76%
Childcare or center-based care	7.67%	16.00%	26.58%	29.84%
Family day care	5.33%	5.33%	9.18%	6.03%
Early Head Start/Head Start	0.33%	4.67%	6.65%	19.05%
Organized play group	4.67%	2.67%	8.86%	8.25%

Child's needs

Table F.15. IEP

	N	Percent	Total
IEP or IFSP?	16	6.32%	253

Food fragility

Table F.16. Food fragility as measured by parental reports on affording meals

Food Fragility	Never	Sometimes	Often	Total
We worried food would run out	72.80%	23.50%	3.70%	243
The food we bought just didn't last	81.90%	14.80%	3.30%	243
We couldn't afford balanced meals	80.50%	17.40%	2.10%	241
We relied on only a few kinds of low-cost meals	78.30%	17.10%	4.60%	240