

EVALUATION OF EARLY  
CHILDHOOD PROGRAMS & CHILD  
DEVELOPMENT IN INDIANA  
Second report  
Commissioned by Early Learning Indiana

---

May 2023

*Milagros Nores,  
Ph.D.,  
Erin Harmeyer,  
Ph.D.,  
Lori Connors-Tadros,  
Ph.D., Zijia Li, Ph.D.,  
Carol Contreras,  
M.Ed.*

*The National  
Institute for Early  
Education Research*

## About the Authors

**Milagros Nores, Ph.D.** is Co-Director of Research and Associate Research Professor and Associate Research Professor 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.

**Erin Harmeyer, Ph.D.** is Assistant Research Professor at NIEER. Dr. Harmeyer conducts research at NIEER related to early childhood programs and evaluation, supporting projects in West Virginia, Indiana, New Jersey, and Pennsylvania.

**Lori Connors-Tadros, Ph.D.** is Senior Research Fellow at NIEER. Dr. Connors-Tadros conducts research to inform state policy makers on issues related to the quality and sustainability of programs for young children birth through early elementary. She is the executive vice president of the Board of the National Parents as Teachers organization.

**Zijia Li, Ph.D.** is Assistant Research Professor at NIEER. Dr. Li is an experienced psychometrician and statistician. She has led and participated leading and conducting rigorous reliability and validity research studies for multiple subjects, including Peabody Picture Vocabulary Test III and IV, the Hawaii Early Learning Profile®, and the High/Scope Child Observation Record.

**Carol Contreras, M. Ed.,** is Program Coordinator at NIEER. She has supported research in West Virginia, Colombia (South America), Indiana, and New Jersey.

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

Permission is granted to reprint this material if you acknowledge NIEER and the authors. For more information, call the Communications contact at (848) 932-4350, or visit NIEER at [nieer.org](http://nieer.org).

Suggested citation: Nores, M., Harmeyer, E., Connors-Tadros, L., Li, Z., & Contreras, C. (2023) Evaluation of the Early Childhood Programs and Child Development in Indiana. Report 2. New Brunswick, NJ: National Institute for Early Education Research.

## Table of Contents

Table of Contents .....	ii
Summary .....	1
Study Methods .....	2
1. Sample.....	3
2. Measures and Procedures.....	4
<i>Classroom Observations</i> .....	4
<i>Teacher self-report quality</i> .....	5
<i>Child Measures</i> .....	6
<i>Parent Survey</i> .....	7
Results.....	7
1. Classroom Observations .....	7
<i>CLASS Results</i> .....	7
<i>CLASS Domains for selected center characteristics</i> .....	13
<i>CLASS comparison to other programs</i> .....	15
2. Teacher characteristics & self-reported quality .....	17
<i>Teacher socio-demographics, work experience, wages, and benefits</i> .....	17
<i>Professional Development Opportunities</i> .....	18
<i>TSEEQ</i> .....	19
<i>TSEEQ Subscale Scores</i> .....	20
<i>TSEEQ Subscale Scores for selected center characteristics</i> .....	21
<i>DEC Checklists</i> .....	22
3. Children’s development in Indiana’s ECE programs .....	25
<i>Infants &amp; Toddlers</i> .....	25
<i>Preschool Age Children</i> .....	29
4. Parent surveys .....	35
<i>Resources and activities in the home</i> .....	36
<i>Parental preferences on child care or preschool program</i> .....	36
<i>Parental perceptions on teachers’ support to their children</i> .....	37
<i>Parental perceptions on what constitutes a high quality early child care and education</i> ....	38
Discussion of Findings.....	43
Acknowledgments.....	46
References .....	47
Appendix A. Measures.....	49

Appendix B. Findings on Classroom Observations..... 52  
Appendix C. Findings on Baseline Child Assessments ..... 54  
Appendix D. Parent surveys ..... 71

## Summary

The National Institute for Early Education Research (NIEER) conducted a landscape evaluation of early childhood programs in Indiana (IN) between the spring of 2021 and the summer of 2022. The evaluation focused on understanding program components, quality, and children’s learning and development across a variety of programs in the state. Commissioned by Early Learning Indiana (ELI), this study aims to provide Indiana programs and policymakers with research-based information on the quality of early childhood programs and the learning and development status of young children birth to five in the state.

The project assesses the landscape of quality in early care and education programs serving children birth to kindergarten in IN, assesses the gains of children enrolled in these programs in language, cognitive, and social-emotional domains, and describes differences observed and reported in program quality by program type, Paths to Quality (PTQ) star level, and geographic region. This second report summarizes the data on program quality across two time points, teacher’ and parents’ perceptions, and the developmental gains of children in sample programs. Information was collected between March 2021 through July 2022. Through this period, programs continued to experience interruptions or quarantines due to the COVID-19 pandemic (including the omicron and delta variants of COVID-19).

This report summarizes classroom quality experienced by children in a sample of 321 classrooms in 206 programs in in Indiana. In addition, the report describes developmental growth in the sampled programs for infants, toddlers, and preschoolers. Findings demonstrate that across 2021 and 2022, infant, toddler, and pre-K classrooms were nurturing and caring environments. Classrooms were observed to be on average moderate to high quality in responsive caregiving (infants), emotional behavioral support (toddlers), and emotional support and classroom organization (preschool) domains. In contrast, areas related to language and instructional support demonstrated low scores across all types of classrooms. This resembles patterns in other similar studies of early care and education programs across the country. We explored quality separately for some subgroups of interest - for (PTQ) level, facility type (auspice), and urban classification. Small differences emerged. Notably, PTQ 3- and 4-rated programs tended to score higher than other programs on infant and toddler observed quality (measured with the Classroom Assessment Scoring System - CLASS instruments) and in instructional support on the pre-K CLASS instrument. On average, infant and toddler classroom quality as measured by the CLASS was found to be consistently higher than those in Early Head Start classrooms,<sup>1</sup> while preschool quality was found to be in the mid-to-low range relative to other state and city programs in the U.S.

In addition, through teacher surveys, we explored teacher self-reported quality using two instruments and teachers’ work experiences and professional development opportunities. Patterns for self-reported quality mostly resemble findings from the observed quality measure, with socio-emotional interactions scoring higher. In addition, teachers reported lower levels of quality in relation to the physical environment. In relation to the teacher self-report checklists, about half of the teachers reported doing the various activities related to informed clinical reasoning most of

---

<sup>1</sup> As in the Baby FACES study, which is discussed later.

the time. This was more frequent in local education agency (LEA) classrooms and centers rated 4. Moreover, teachers report higher frequency in the use of assessments to inform their work with children, with no differences emerging across program types, rating, or urbanicity.

Descriptions on the developmental growth of children in early childhood programs provide a snapshot on the development of children in several types of ECE programs in Indiana. We find lower developmental levels in language and cognition for infants and toddlers, and socio-emotional levels at par with what we expect due to maturation. However, we also find growth across infants and toddlers higher than typical for this group age. At baseline, the children of preschool age in the sample had lower levels of literacy and math skills compared to their typical peers, but they had similar language and executive function skills. During the study period, the children made strong gains in vocabulary, which was higher than typical growth for children their age, and showed expected growth in math and executive functions. However, their literacy skills showed slower growth than typical, and in fact, they had a loss relative to their peers in standard scores. The patterns were consistent across ratings, urbanicity, and program type, with all children showing significant growth in receptive vocabulary, regardless of their enrollment location or center type. However, literacy development lagged for all children similarly, regardless of where children were enrolled. Overall, children showed improvements in externalizing behaviors.

Finally, in relation to parental perceptions, we found that parents reported frequently engaging in activities that support children's learning, with the majority reporting that they do things like talk with their child about the world around them daily. Most reported feeling noticeably confident that they fully understand what constitutes a high-quality environment, and having providers who are warm and caring emerged as the most important aspect of a quality environment to them, both in quantitative and qualitative data. These perceptions of quality and the environment children were learning in differed somewhat for parents of pre-K when compared to infant/toddler-aged children, and these differences are explored further in the report.

## Study Methods

This evaluation of Indiana's ECE system is a multi-site study encompassing several components to provide the first-ever independent assessment at this scale of the quality of early care and education programs of various auspices and to assess the developmental growth of children in those programs. We present the findings of independent assessments on multiple dimensions of program quality and selected program characteristics. Data collection included administrator surveys and program observations conducted between the spring and fall of 2021, and again in the spring of 2022. We also report developmental growth in children within this period across programs. This report addresses the following research questions using this information:

1. What is the observed quality of children's classroom experiences across two observations in the study period? Are there any observed changes in quality between 2021 and 2022?
2. What are teachers' self-reported quality on classroom practices?
3. What is the developmental growth of children enrolled in early childhood programs?
4. What are parents' perceptions of programs and program quality?

The study’s original plan was to measure classroom quality in the spring of 2021 and to measure child progress and program quality in the school year 2021-22, but the study had to revise these timelines to adapt to the constraints imposed by the COVID-19 pandemic. Procedures and measures are described in detail below. Classroom observations were conducted starting in May 2021 and through the end of December 2021, then once again starting in March 2022 and through June 2022. Child assessments were conducted twice to assess child growth over time - once between September 2021 and February 2022 and again between April 2022 and July 2022. Measures and procedures for the evaluation are summarized further below. The delays and constraints imposed by COVID-19 and the delta and omicron variants on early childhood programs required working with programs to engage them in the study and to find a time for observations and child assessments in which the program felt they could be visited by the research team. It is important to recognize that survey data was the most impacted by this, and this included surveys that asked teachers to report on the socio-emotional status of the children.

## 1. Sample

The sample is composed of 321 classrooms in 206 programs (48 of which were home-based providers and 29 of which were affiliated with Building Blocks, a non-profit serving 28 counties in Southern Indiana). The first report (Nores, et al., 2022) described program recruitment and observation procedures. Classroom quality was observed twice in each classroom using the CLASS Infant, Toddler, or Pre-K depending on classroom composition.<sup>2</sup> Table 1 below reports baseline program characteristics for the recruited classroom sample across types of programs, quality ratings, and county classification.

Table 1. Sample characteristics for programs.

Program Characteristics		Count	%
<b>Facility Type</b>	Center Licensed	70	33.99%
	Home Licensed	48	23.30%
	LEA	18	8.74%
	Ministry	58	28.16%
	Other*	11	5.34%
<b>Quality Rating</b>	1	30	14.56%
	2	11	5.47%
	3	84	40.76%
	4	55	26.70%
	Non-Rated**	25	12.14%
<b>County Classification</b>	Mid-sized	46	22.33%
	Rural	71	34.47%
	Urban	89	43.20%
<b>Community Poverty</b>	Low	158	78.61%
	High	43	21.39%

Note: Administrative surveys were completed by 228 centers, including some centers that dropped out of the study.

<sup>2</sup> Building Blocks of Indiana independently completed and provided CLASS scores for ten infant, 24 toddler, and 61 pre-K classrooms. NIEER observed 226 classrooms. These scores are included in this report and in Table 1. Building Blocks programs were only observed once.

\*Programs noted as “other” are those that are not classified as eligible to be included in the Paths to Quality quality rating and improvement system (QRIS) (mostly public and private school-based programs that are legally license-exempt). \*\*Non-rated includes both programs rated 0 and those categorized as “other” and are thus non-rated.

Tables 2 reports the number and percentage of children across selected characteristics, including center characteristics (rating, urbanicity) and community poverty<sup>3</sup> by age group. We report children in the sample for which we have two assessment points in order to measure growth.

Table 2. Baseline children’s sample characteristics

		Infant (n=223)	Pre-K (n=492)
<b>Gender</b>	Female	52.47%	52.64%
	Male	47.53%	47.36%
<b>Age Cohort</b>	1	17.49%	n/a
	2	40.81%	n/a
	3	41.70%	n/a
	4	n/a	30.89%
	5	n/a	53.86%
	6	n/a	12.60%
<b>By Center Rating</b>	1 & 2	24.66%	17.07%
	3	29.60%	43.70%
	4	30.49%	24.39%
	Unrated	15.25%	14.84%
<b>By Urbanicity</b>	Mid-size	23.32%	25.61%
	Rural	32.29%	29.07%
	Urban	44.39%	45.33%
<b>By Community Poverty</b>	Low	19.19%	22.59%
	High	80.81%	77.41%

## 2. Measures and Procedures

### *Classroom observations*

In both 2021 and 2022, classroom quality was measured using CLASS. The study included the use of three versions depending on a classroom’s predominant age group: Infant (CLASS Infant; Hamre et al., 2014), Toddler (CLASS Toddler; LaParo et al., 2012), and Pre-K (CLASS Pre-K; Pianta, La Paro, & Hamre, 2008). The CLASS captures teacher-child interactions and classroom processes.<sup>4</sup> More detail on the CLASS is provided in Appendix A. Observers were trained in

<sup>3</sup> We have adopted the definition of poverty areas based on the U.S. Census Bureau’s 2015-2019 5-year American Community Survey (ACS), which defined as census tracts where at least 20% of the population lives in poverty. Source: <https://www.census.gov/library/stories/2021/02/fewer-people-living-in-poverty-areas-2015-2019.html>. The data on poverty is from the Social Vulnerability Index data: <https://www.atsdr.cdc.gov/placeandhealth/svi/>.

<sup>4</sup> The protocol used required that at least four children were present, and at least half of the children’s ages aligned with the CLASS tool used (e.g., at least half of the children were ages birth – 18 months to be observed with CLASS Infant). In classrooms with an approximately even split of children (e.g., an FCC home with three pre-K children



reliability by a CLASS Affiliate Trainer, or through the Teachstone® training platforms. All observers met the developer’s reliability requirements (80%) for observer certification. Observers were also trained in practices and procedures for conduct and required to complete background checks and training in human subjects’ research (human subject protections, ethical issues, etc.). All assessors were required to pass a calibration mid-data collection.

The previous report (Nores, et. al, 2022) discussed the findings for paired environmental checklists and a program survey. This report focuses on results from a teacher survey deployed throughout the project period to the sample.<sup>5</sup> The teacher survey includes information on self-reported quality and staff qualifications and background.

### *Teacher self-report quality*

The teacher survey included questions about teacher socio-demographic background, education, certification, income and benefits, years teaching and in the current position, and professional development opportunities. In addition, the surveys included the following instruments for self-reported quality:

*TSEEQ: The Teacher Survey of Early Education Quality*<sup>6</sup> is a self-report survey about early childhood classroom practices for quality administered to early childhood educators. The survey is completed independently and can be done on paper or online. Teachers are asked to reflect on aspects of their classroom practice including curriculum, instruction, assessment, leadership and supervision, physical environment, interaction and emotional climate, and family involvement. Within these aspects are questions about several classroom practices, including in the areas of literacy, science, and math. The survey consists of 105 questions, and most of these are answered on a five-point Likert Scale or involve a yes/no response. The survey takes about 20 minutes to complete.<sup>7</sup> Adaptations were made for the infant/toddler age group and for family child care (FCC) providers.

*DEC Recommended Practices (RP) Checklists:* The teacher surveys also included a subset of performance checklists for program improvement.<sup>8</sup> We included the following two which focuses on the assessment of children: (i) The Informed Clinical Reasoning Checklist (ICRC), and (ii) The Authentic Child Assessment Practices Checklist (ACAPC). The ICRC includes practices important for evaluation and eligibility determination. The ACAPC includes characteristics of authentic assessment practices for observing children’s everyday activities and learning opportunities. Both of these use a four-point Likert system going from a practice occurring “seldom or never” to “most of the time.”

---

and three infants), two CLASS tools were used, rotating cycles between the two as recommended by the developer. Given the smaller size of FCCs, we required at least two children present.

<sup>5</sup> Teachers in the Building Blocks programs did not participate in this survey.

<sup>6</sup> Hallam, Rous, Riley-Ayers, & Epstein (2011).

<sup>7</sup> Eleven questions on the TSEEQ were slightly revised for home-based providers to be more applicable to their setting. For example, the question, “I know the evaluation process and tools my supervisor uses to assess my performance” was revised to “I know the evaluation process and tools my supervisor *or coach* uses to assess my performance” for home-based providers.

<sup>8</sup> <https://ectacenter.org/decrp/type-checklists.asp>

### *Child measures*

This study employed a set of standardized child assessments designed to measure learning across various domains and that are psychometrically valid, proven to discriminate effects in intervention studies, and appropriate for the age range of birth to five. Analyses of growth in children in these measures allows for the understanding of the differential experiences of children across programs in Indiana. Children assessed were enrolled in the programs that agreed to participate in this study between the months of September of 2021 and February of 2022 (pretest) and then again between April and July of 2022 (posttest). This evaluation includes infants and toddlers as well as preschool-age children; therefore, assessment measures are split into two separate categories as they differ across these two age groups.

Preschool-age children were assessed in expressive vocabulary, math, literacy, executive function, and socio-emotional skills. The measures utilized are:

- The Peabody Picture Vocabulary Test--Fourth Edition (PPVT-IV; Dunn & Dunn, 2007), a 204-item test of receptive vocabulary in standard English.
- The Woodcock-Johnson Psycho-Educational Battery-Fourth Edition (WJ-IV; Schrank, Mather, & McGrew, 2014), which includes two subscales: Applied Problems and Letter-Word Identification subtests, which broadly measure math and "reading."
- Dimensional Change Card Sort Task (DCCS; Zelazo, 2006) assesses attention-shifting.
- Peg Tapping Test (PT; Diamond & Taylor, 1996). This test requires children to inhibit a natural tendency to mimic the experimenter while remembering the rule for the correct response.
- The Child Behavioral Checklist (Achenbach, 2009), a teacher (or parent) reported measure of children's social-emotional skills.

Infants and toddlers were assessed with the Bayley Scales of Infant and Toddler Development-Fourth Edition (Bayley-IV). This is a comprehensive assessment of five developmental domains for children ages 1 to 42 months of age. The Bayley Scales of Infant Development (BSID) are the most commonly used assessment of infant development (Fernald, Kariger, Engle, and Raikes, 2009). We used two scales of the Bayley —cognitive and language (expressive and receptive), and the socio-emotional accompanying surveys to teachers. The Bayley scales have been shown to predict later non-verbal and verbal cognition and have been extensively validated on other measures (Blaga, et. al, 2009; Feinstein, 2003).

Preschool child assessors were trained to reliability by the NIEER team, and Bayley IV assessors were trained through Pearson's online training platforms, with reliability conducted with NIEER. NIEER Training includes training in practices and procedures for research conducted with children, as well as completing a background check and training in human subjects' research (human subject protections, ethical issues, etc.).

### *Parent survey*

Parents of assessed children were provided with a survey that could be completed on paper or online. The parent survey included 17 questions (many with multiple sub-questions) that address their perceptions of their child’s early care and education program, their beliefs on what best represents quality for young children, and their confidence in their understanding of what constitutes a high-quality environment. Socio-demographic questions were also included as part of the survey.

## **Results**

Results are reported first for the CLASS, and in relation to programs’ PTQ ratings, their urban, mid-size, and rural county location, and their program type. The second section describes self-reported teaching and evaluation practices by teachers. We then describe the findings on developmental gains for children in the programs by age, urbanicity, program rating, and program type, then estimates of the association between children’s gains and various center and family variables. The last section reports parent surveys. The report concludes with a discussion of the findings.

### **1. Classroom Observations**

#### *CLASS results*

Average CLASS scores for the sample of Indiana early childhood classrooms for all domains and dimensions are reported in Table 3 by age group and for the two time points and the average across these. Patterns of observed quality are consistent of typical patterns in the field for other studies of state early childhood education programs (discussed further below).

- CLASS Infant ratings for programs observed showed an average rating of 4.84 for responsive caregiving (RC), with a minimum score of 2.42 and a maximum score of 6.43.
- Classrooms rated with the CLASS Toddler were an average 5.69 for Emotional and Behavior Support (EBS), and 3.25 for Engaged Support for Learning (ESL), with minimum scores observed at 2.95 and 1.00 and maximum scores at 6.76 and 5.87, respectively.
- CLASS Pre-K scores were rated on average 5.80 for Emotional Support (ES) (minimum of 2.58 and maximum of 7.00), 5.16 for Classroom Organization (CO) (minimum of 1.44 and maximum of 6.73), and 2.78 for Instructional Support (IS) (minimum of 1.00 and maximum of 5.33).

Across the different age groups, results point to lower-than-average scores on the sub-components and domains that relate to supports for language, learning, and overall instruction. Specifically, for infants, “facilitated exploration” and “early language supports” scored lowest. Similarly, in toddler classrooms, all the domains in “engaged supports for learning” scored lower. For preschoolers, all the domains in “instructional support” scored lower.

Table 3. CLASS Domain and Dimension means and ranges, by age group and data collection point

Domain & Dimensions	2021			2022			Average		
	Mean	Min.	Max.	Mean	Min.	Max.	Mean	Min.	Max.
<b>Infant (n=39)</b>									
<b>Responsive Caregiving (RC)</b>	<b>4.99</b>	<b>2.42</b>	<b>6.70</b>	<b>4.78</b>	<b>2.17</b>	<b>6.65</b>	<b>4.84</b>	<b>2.42</b>	<b>6.43</b>
Relational Climate (RL)	5.71	3.33	7.00	5.34	2.33	7.00	5.52	3.20	7.00
Teacher Sensitivity (TS)	5.58	3.20	7.00	5.30	2.33	7.00	5.40	3.20	7.00
Facilitated Exploration (FE)	4.61	1.00	6.60	4.44	1.67	6.20	4.48	1.00	6.20
Early Language Support (ELS)	4.06	1.67	7.00	4.03	1.80	6.40	3.95	1.67	6.50
<b>Toddler (n=92)</b>									
<b>Emotional and Behavior Support (EBS)<sup>a</sup></b>	<b>5.78</b>	<b>3.05</b>	<b>6.88</b>	<b>5.50</b>	<b>2.67</b>	<b>6.72</b>	<b>5.69</b>	<b>2.95</b>	<b>6.76</b>
Positive Climate (PC)	6.00	2.75	7.00	5.36	2.33	7.00	5.81	2.77	7.00
Negative Climate (NC)	6.87	5.80	7.00	6.85	5.80	7.00	6.87	6.00	7.00
Teacher Sensitivity (TS)	5.64	2.25	7.00	5.17	1.67	6.80	5.47	2.03	7.00
Regard for Child Perspectives (RCP) <sup>a</sup>	4.95	1.75	7.00	4.92	1.67	6.60	4.97	1.73	6.60
Behavior Guidance (BG)	5.44	2.20	7.00	5.18	1.67	7.00	5.34	1.93	7.00
<b>Engaged Support for Learning (ESL)</b>	<b>3.07</b>	<b>1.00</b>	<b>5.67</b>	<b>3.48</b>	<b>1.00</b>	<b>5.87</b>	<b>3.25</b>	<b>1.00</b>	<b>5.87</b>
Facilitation of Learning and Development (FLD)	3.59	1.00	6.80	3.69	1.00	6.40	3.65	1.00	6.40
Quality of Feedback (WF)	2.61	1.00	5.40	3.03	1.00	6.00	2.80	1.00	6.00
Language Modeling (LM)	3.00	1.00	5.80	3.72	1.00	6.60	3.30	1.00	6.40
<b>Pre-K (n=203)**</b>									
<b>Emotional Support</b>	<b>5.74</b>	<b>3.60</b>	<b>7.00</b>	<b>5.86</b>	<b>2.58</b>	<b>7.00</b>	<b>5.80</b>	<b>2.58</b>	<b>7.00</b>
Positive Climate (PC)	5.97	3.20	7.00	5.87	2.00	7.00	5.93	2.00	7.00
Negative Climate (NC)	6.76	1.40	7.00	6.83	5.40	7.00	6.80	4.00	7.00
Teacher Sensitivity (TS)	5.40	2.40	7.00	5.64	1.00	7.00	5.52	1.00	7.00
Regard for Student Perspectives (RSP)	4.84	1.60	7.00	5.09	1.67	7.00	4.93	1.67	7.00
<b>Classroom Organization (CO)</b>	<b>5.08</b>	<b>2.13</b>	<b>6.73</b>	<b>5.26</b>	<b>1.44</b>	<b>6.87</b>	<b>5.16</b>	<b>1.44</b>	<b>6.73</b>
Behavior Management (BM)	5.46	1.60	7.00	5.62	2.00	7.00	5.53	2.00	7.00
Productivity (PR)	5.40	2.20	7.00	5.57	1.33	7.00	5.48	1.33	7.00
Instructional Learning Formats (ILF)	4.37	1.40	6.60	4.60	1.00	7.00	4.47	1.00	6.40
<b>Instructional Support (IS)</b>	<b>2.76</b>	<b>1.00</b>	<b>5.92</b>	<b>2.75</b>	<b>1.00</b>	<b>5.92</b>	<b>2.78</b>	<b>1.00</b>	<b>5.33</b>
Concept Development (CD)	2.62	1.00	5.75	2.54	1.00	6.00	2.62	1.00	5.50
Quality of Feedback (QF)	2.88	1.00	6.00	2.87	1.00	6.25	2.90	1.00	5.75
Language Modeling (LM)	2.77	1.00	6.00	2.85	1.00	6.00	2.83	1.00	5.13

\*The Negative Climate dimension is reverse scored so that a high score represents “good.”

\*\*Classrooms scored with multiple tools have both scores reflected in this table.

<sup>a</sup>. Statistically significant differences between 2021 and 2022 distribution of scores were assessed. Only the toddler domains were significantly different at a 5% level.

Having assessed the classroom quality twice in each classroom, we are also able to assess overall stability in average domain scores across the sample. Given the impact that the COVID-19 pandemic was still having through 2021 on child care and education services across the nation (Weiland et al., 2021), a second data collection point in 2022 allowed for understanding the degree to which areas of concern flagged in the first report (Nores, et. al, 2022) may have been due to the pandemic, or whether these continued to be of particular concern about a year later. Figure 1 illustrates some degree of stability of scores across these two points in time for both the infant and toddler measures. Figure 2 shows a similar stability in the CLASS Pre-K measure. Only the toddler domains were statistically different between the years, with EBS being lower in the second round and ESL being higher in the second round.

Figure 1. CLASS Infant and Toddler average scores in 2021 and 2022

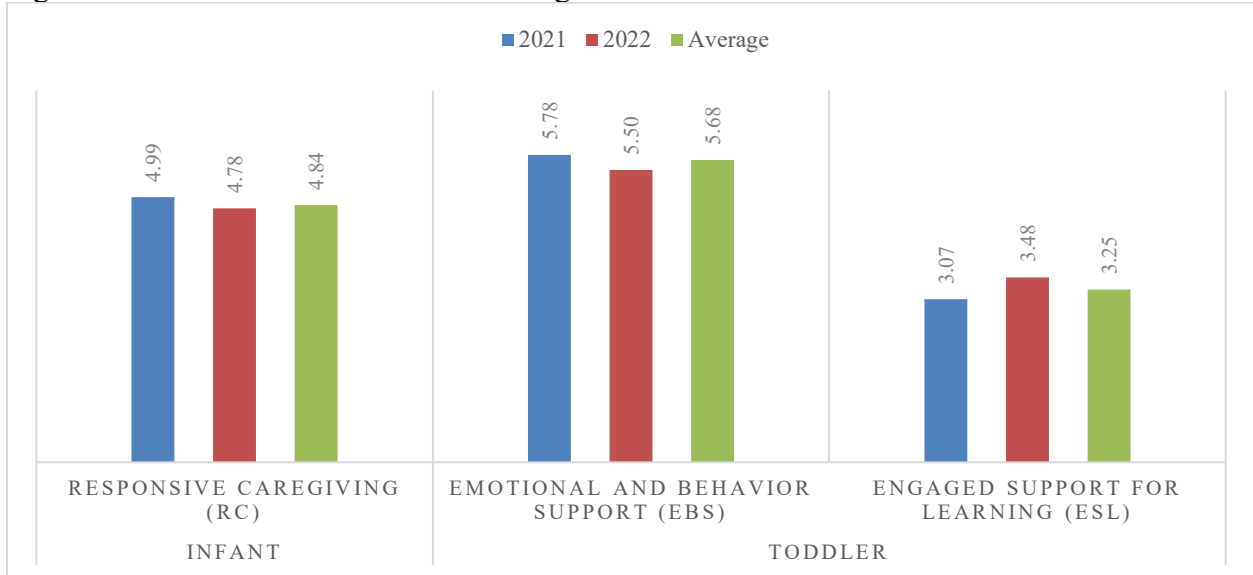
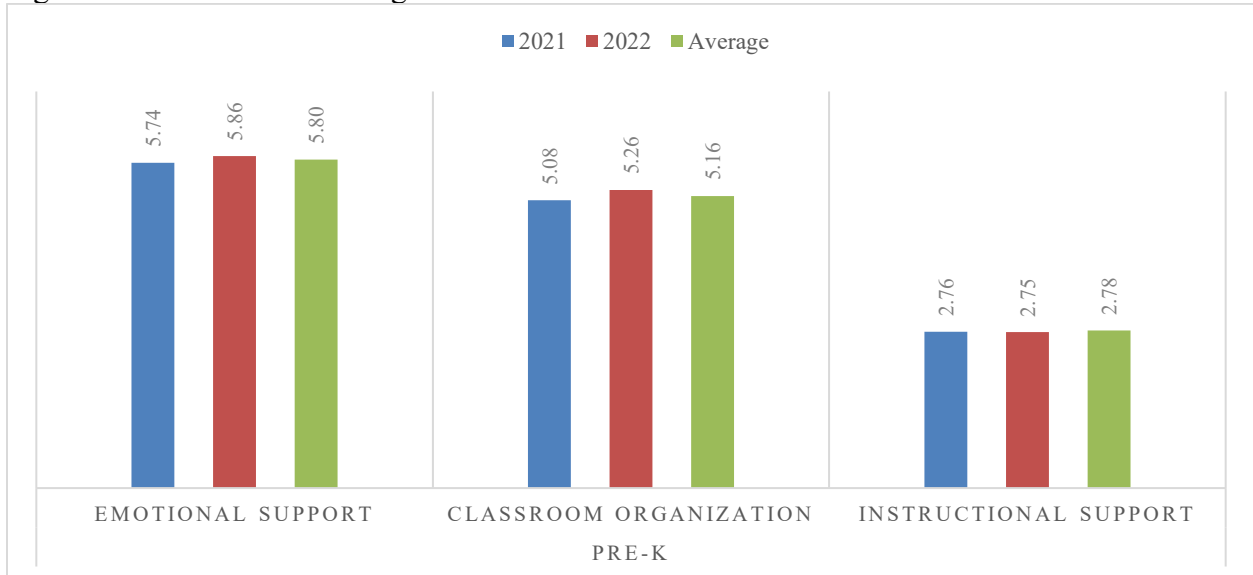


Figure 2. CLASS Pre-K average scores in 2021 and 2022



Overall findings for the CLASS Infant score, the CLASS Toddler EBS domain, and the CLASS Pre-K ES domain on climate and teacher sensitivity provide indication that the classrooms observed are in general caring and warm environments. Some low-scoring classrooms are present in the sample that could benefit from directly addressing these aspects of child care and education processes. On average, these dimensions across the three age-related tools were rated close to or above 5. Aspects of classroom organization and group management are embedded in the CLASS Infant dimension of facilitated exploration and the CLASS Toddler dimensions of Behavior Guidance and Facilitation for Learning & Development. These were rated on average at 4.48, 5.34, and 3.65, respectively. Dimensions related to language support, modeling, productivity, concept development, scaffolding, and feedback to children were rated on average

at levels 3 or under. The ratings show classrooms observed with scores at inadequate levels of quality for aspects related to classroom management and facilitation, as well as language and child scaffolding. This is illustrated in Figures 1 and 2, which show the distributions for each dimension.

Figures 3-8 below demonstrate the variations in scores between infant, toddler, and pre-K measures (across the two time points). By analyzing these distributions, it is possible to gain a more comprehensive understanding of the percentage of classrooms that have scored at low, adequate, and high levels, rather than just focusing on the average score. For example, in preschool classrooms, CLASS scores below 3 purportedly indicate low levels of Emotional Support (ES) and Classroom Organization (CO), whereas scores between 3 and 5 are considered adequate, and scores above 5 or 7 are deemed good or excellent. However, it is essential to consider that some domains have shown lower scores overall across many studies and research indicates that “good” scores on these domains may realistically be above 3.5 or 4. Instructional Support (IS) in preschool classrooms is one such example.

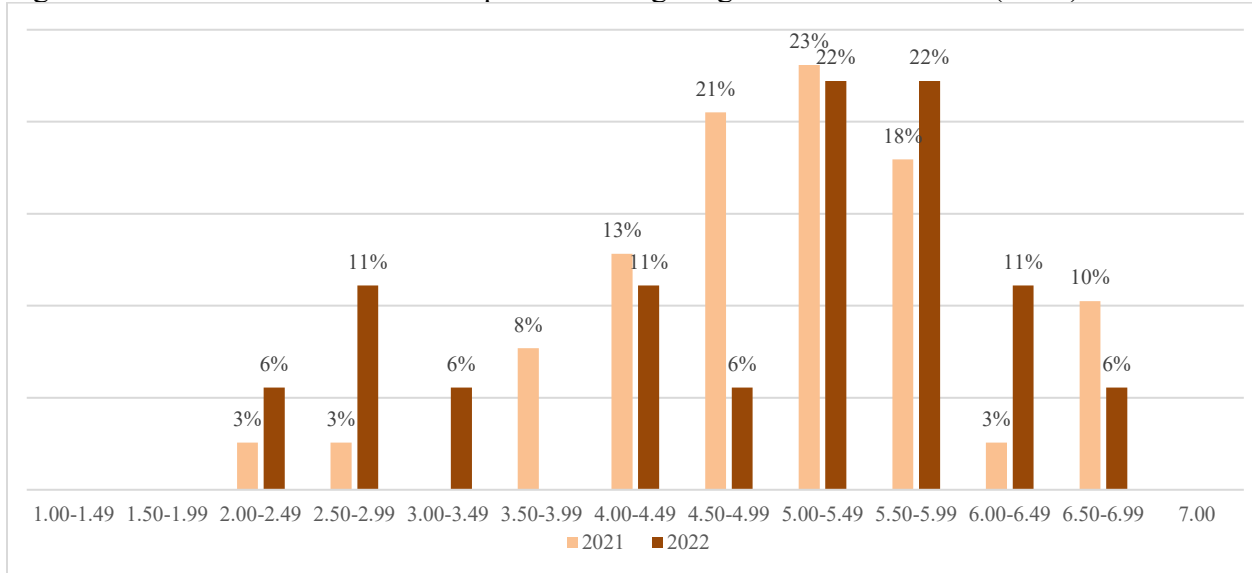
In regards to CLASS pre-K, there is research that suggests a relationship between quality and children's outcomes in pre-K when the scores for CO are above 5 and for IS are above 3 (other studies define these thresholds slightly higher at 5.5 and 3.5) (Burchinal et al., 2009; Burchinal et al., 2014; Hatfield et al., 2016).<sup>9</sup> In this study, we refer to this research to determine the percentage of classrooms that fall above these defined quality thresholds. The key takeaways from these assessments are as follows:

- For infants, 47% of classrooms were rated at or above 5 in Responsive Caregiving.
- For toddlers, 85% of classrooms scored above 5 in Emotional and Behavioral Support, and 54% of classrooms were rated at or above 3 in Engaged Support for Learning.
- For preschool-aged children, 85% of classrooms were rated at or above 5 on Emotional Support, and 61% were rated at or above 5 on Classroom Organization.
- However, just 38% of classrooms scored above 3 on Instructional Support for preschool-aged children.

---

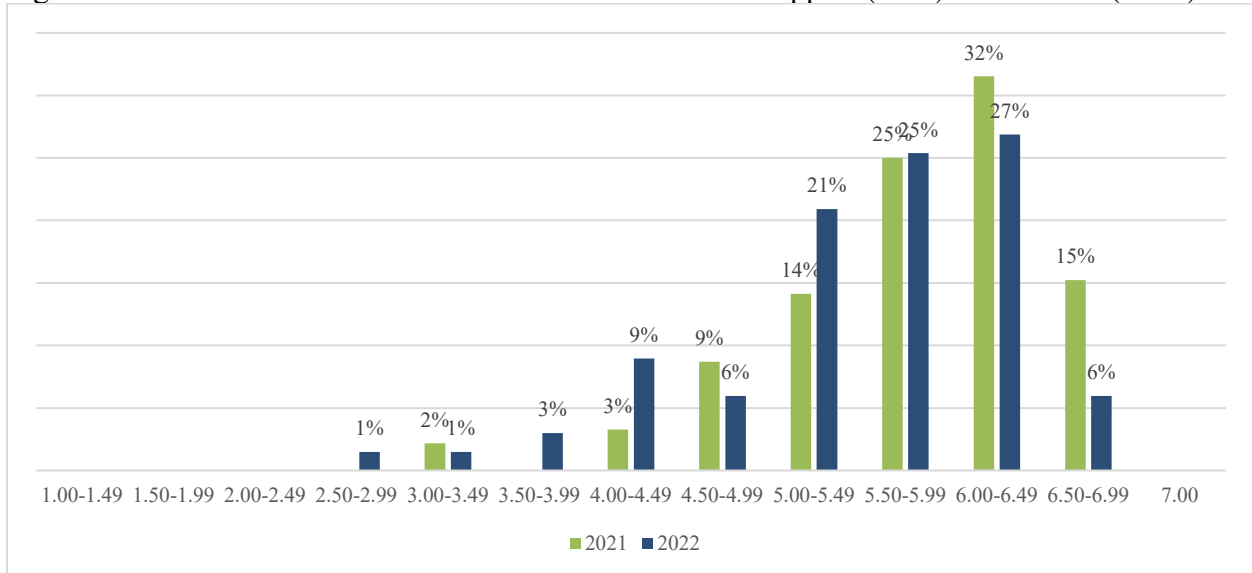
<sup>9</sup> The Office of Head Start defines quality thresholds for CLASS Pre-K at 6 for the Emotional Support domain and the Classroom Organization domain, and at 3 for the Instructional Support domain. See: <https://eclkc.ohs.acf.hhs.gov/designation-renewal-system/article/use-classroom-assessment-scoring-system-class-head-start>. While more research is needed in relation to thresholds, we use these and the cited literature to describe what was observed in the sample.

Figure 3. Distribution of CLASS Responsive Caregiving Domain for infants (n=43).



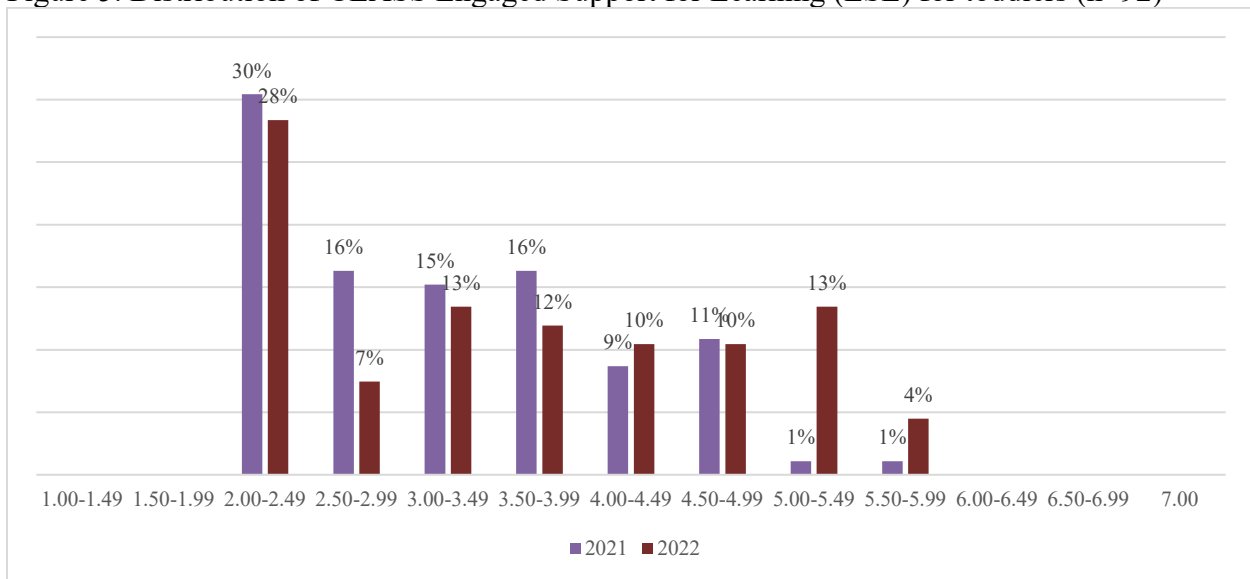
Note: For infants, 54% and 61% of classrooms scored at or above 5 in Responsive Caregiving. A small portion were at the inadequate level (<3), but the majority of the classrooms are at or above 5 (good levels of quality).

Figure 4. Distribution of CLASS Emotional and Behavior Support (EBS) for toddlers (n=92)



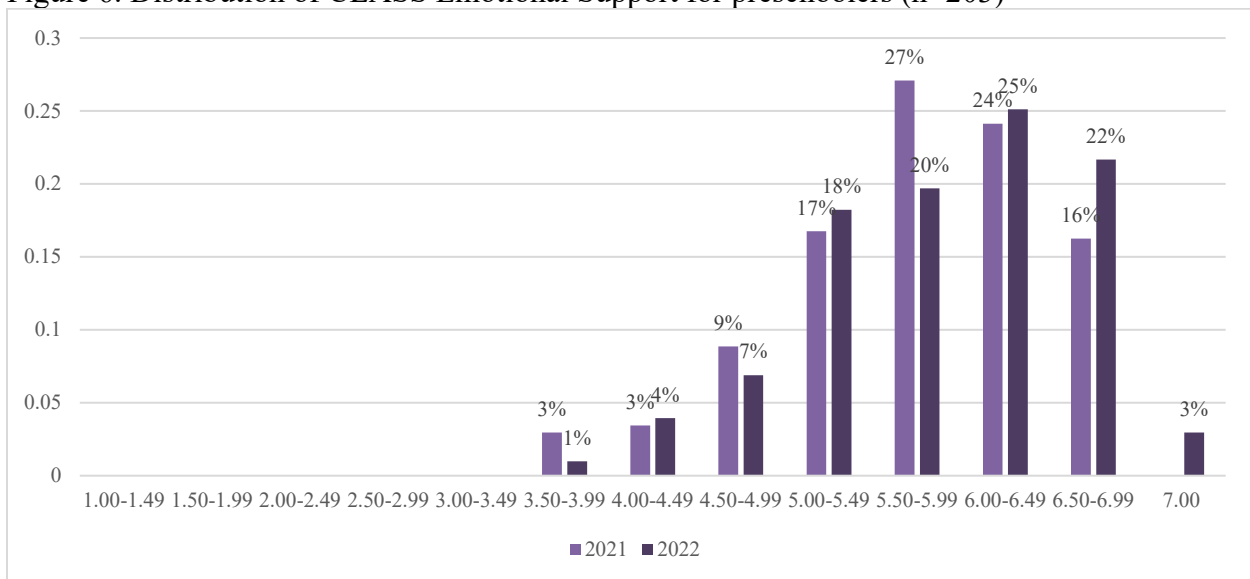
Note: For toddlers, most classroom evidence warm and caring environments with no classroom under 3 (inadequate) for the second observation.

Figure 5. Distribution of CLASS Engaged Support for Learning (ESL) for toddlers (n=92)



Note: For toddlers, between 46% and 35% of the classroom were rated under 3 (inadequate) in the two rounds of data collection. About half to two-thirds were rated above 5, evidencing good levels of quality as per the instrument.

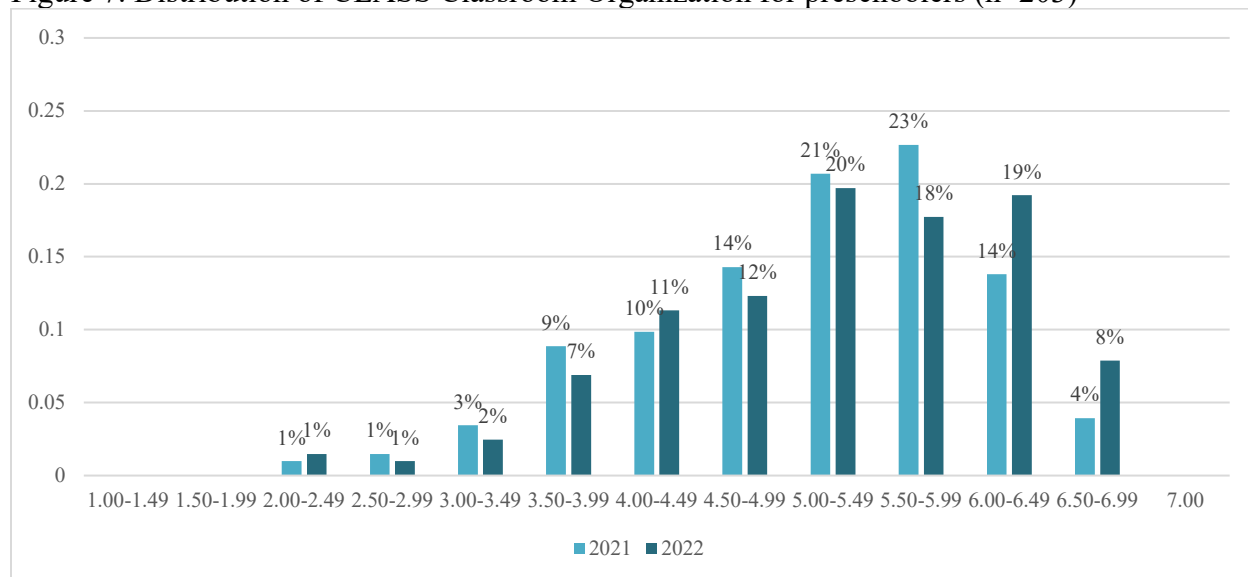
Figure 6. Distribution of CLASS Emotional Support for preschoolers (n=203)



Note: For classrooms serving preschool-age children, there were no classrooms with inadequate levels of emotional support. Over 85% of the classrooms were rated above 5.

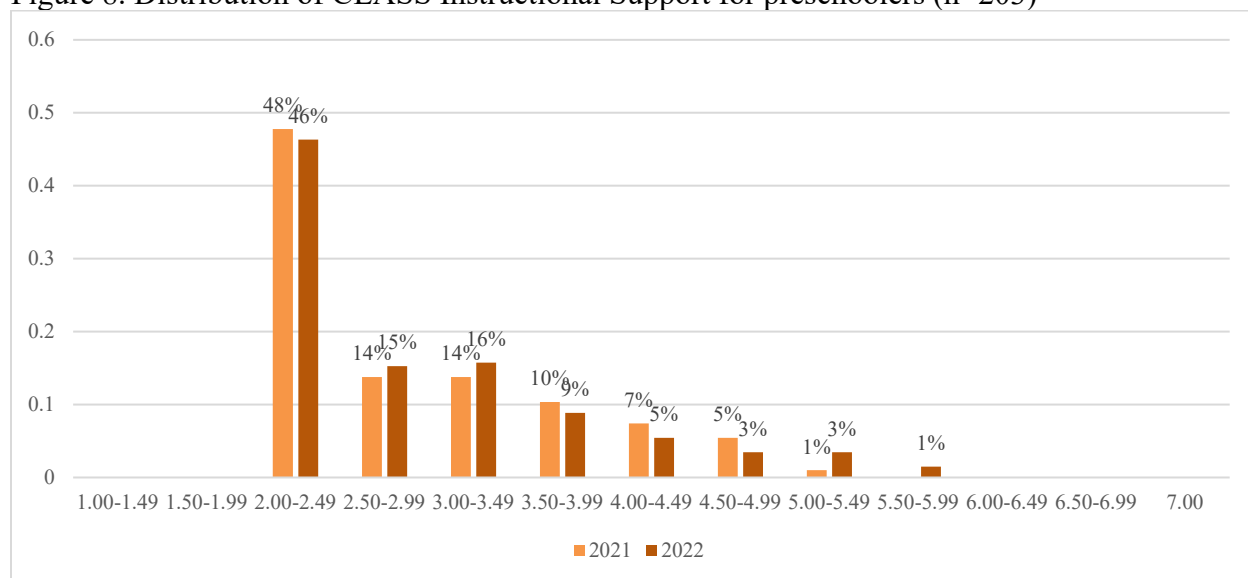


Figure 7. Distribution of CLASS Classroom Organization for preschoolers (n=203)



Note: For classrooms serving preschool-age children, there were few classrooms that rated at the inadequate level of classroom organization. Over 60% of the classrooms were rated above 5.

Figure 8. Distribution of CLASS Instructional Support for preschoolers (n=203)



Note: For classrooms serving preschool-age children, about two-thirds of the classrooms were rated under 3. In both data collection points, 38% of the classrooms were rated at or above 3.

### CLASS Domains for selected center characteristics

We explored differences in CLASS scores by selected program-level characteristics - by program type, PTQ ratings, and urbanicity. Figures are included in Appendix B. Key findings include:

- CLASS scores in infant classrooms differed slightly by program type and PTQ rating, with licensed home-based programs (4.8) and particularly ministry programs (5.2)

scoring slightly higher than center-based (4.7) and other (3.8)<sup>10</sup> programs across 2021 and 2022. Additionally, classrooms with PTQ ratings of 1 and 2 tended to score lower (4.2) on average on CLASS Infant relative to those with PTQ ratings 3 or 4 (5.1 and 4.9, respectively).

- In relation to toddler classrooms, all program types scored similarly on EBS, but ministry programs scored slightly higher on ESL. Classrooms with ratings of 1 and 2 tended to score lower in the EBS and ESL domains of the CLASS relative to centers with PTQ ratings of 3 and 4. However, these were not as different from classrooms in non-rated centers on the ESL domain. Meanwhile, classrooms rated 4 scored the highest in the ESL domain. Finally, programs located in urban areas tended to score lower in both domains of the CLASS Toddler.
- In the CLASS Pre-K, LEA classrooms scored higher in all domains compared to center-licensed and ministry-based programs, especially in the Classroom Organization domain. Regarding the PTQ ratings, most programs scored similarly in the Emotional Support and Classroom Organization domains. However, higher-rated programs (PTQ ratings of 3 and 4) scored higher in the Instructional Support domain, although their average score was still below 3. Differences by urbanicity were minimal.

### *Multivariate analyses*

We conducted multivariate estimates to examine the significance of these differences in CLASS scores across various program types and settings. Results showed that for pre-K classrooms, LEA programs had significantly higher scores across all domains of the CLASS tool, with this effect being sustained in the Classroom Organization domain even after controlling for program rating, urbanicity, and poverty. Home-based licensed programs had significantly lower scores in the Instructional Support domain. In contrast, programs rated 4 had significantly higher scores in both Toddler scales, and this effect was observed in Pre-K Emotional Support and Instructional Support even after controlling for program type and urbanicity. Lastly, programs rated 1 and 2 had significantly lower scores in the Instructional Support domain.

---

<sup>10</sup> Programs noted as “other” are those that are not classified by the Paths to Quality QRIS (typically school-based programs that are legally license-exempt).

Table 4. CLASS 2021-2022 average domain mean scores by subgroups.

	INFANT (n=39)	TODDLER (n=92)		Emotional Support	PRE-K (n=203)	
	Responsive Caregiving	Emotional and Behavioral Support	Engaged Support for Learning		Classroom Organization	Instructional Support
<b>Facility Type</b>						
Center Licensed	4.7	5.7	3.2	5.8	5.1	2.8
Home Licensed	4.8	5.5	3.0	5.6	5.0	2.5
LEA	.	.	.	6.1	5.8	3.2
Ministry	5.2	5.8	3.5	5.8	5.1	2.7
Other*	3.8	5.6	3.2	6.0	5.3	3.1
<b>Quality Level</b>						
1 & 2	4.2	5.3	2.9	5.6	4.9	2.2
3	5.1	5.7	3.3	5.8	5.2	2.9
4	4.9	6.0	3.7	5.9	5.3	2.8
Non-rated	4.9	5.6	2.8	5.8	5.2	2.7
<b>County Classification</b>						
Mid-sized	4.6	5.7	3.5	5.9	5.2	2.7
Rural	4.7	5.7	3.3	5.8	5.2	2.9
Urban	5.0	5.7	3.1	5.8	5.1	2.7

\*Programs noted as “other” are those that are not classified by the Paths to Quality QRIS (school-based programs that are legally license-exempt). Note: Two-tailed tests of differences in means were run between all groupings for each domain. Statistically significant differences were only found between LEA and Center-licensed, LEA and Ministry pre-K classrooms in the Classroom Organization domain, and LEA and Center-licensed pre-K classrooms in the Instructional Support domain. The range of classrooms for each subgroup reported in the table above was from 4 to 92, except for the category of “Other” under facility type for infants for which only one classroom (an infant/toddler program located in and run by a school district) was captured and therefore not reported.

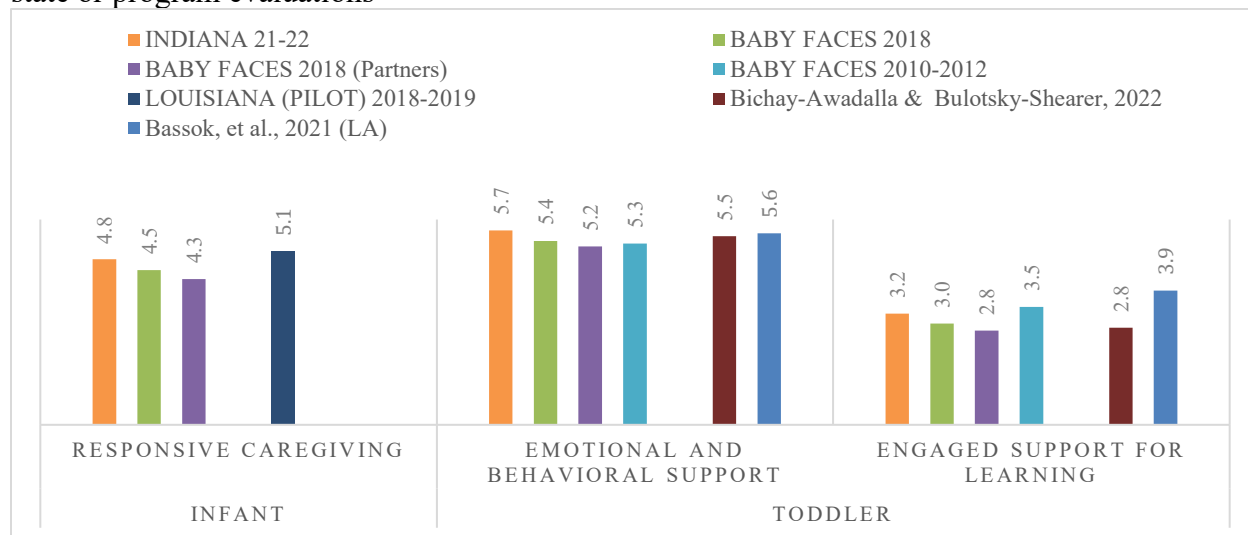
*CLASS comparison to other programs*

To provide context for the quality of infant and toddler classrooms observed for the sample in this study, we compare our findings to those of other relevant infant and toddler studies in the field. The CLASS Infant and Toddler assessment tools have not been as commonly used in the field of early childhood care and education as the CLASS Pre-K tool, resulting in a limited number of representative state or city evaluations available for comparison. To provide a benchmark for interpreting scores in Indiana, we refer to a recent report from the Early Head Start Family and Child Experiences Survey (Baby FACES; Xue, 2021), in which the authors report an average score of 4.52 (with a minimum of 2.31 and a maximum of 6.69) on the Responsive Caregiving domain across 149 classrooms observed using the CLASS Infant tool. The specific dimensions within the Responsive Caregiving domain scored 5.39 on average for Relational Climate, 5.18 for Teacher Sensitivity, 3.84 for Facilitated Exploration, and 3.64 for Early Language Support. Comparatively, the scores obtained using the CLASS Infant tool in the current study were higher than those summarized in Baby FACES across all four dimensions.

Likewise, the Baby FACES study assessed a sample of 713 toddler classrooms using the CLASS Toddler tool, with average scores of 5.37 in Emotional and Behavioral Support and 2.96 in Engaged Support for Learning. In contrast, the observed scores for the Indiana sample in this study surpassed those of the Baby FACES study for all domains and dimensions in the CLASS

Toddler assessment, except for Language Modeling. These comparisons are illustrated in Figure 9. It is important to note, however, that the samples from the Baby FACES study and our study differ in terms of program type and geographic location. However, these comparisons provide a benchmark for interpreting the results of this study and suggest that the infant and toddler classrooms observed are quite at par with what has been observed in other programs, although mostly at or below 5 for infants and only seem stronger for the emotional support aspects in toddler classrooms.

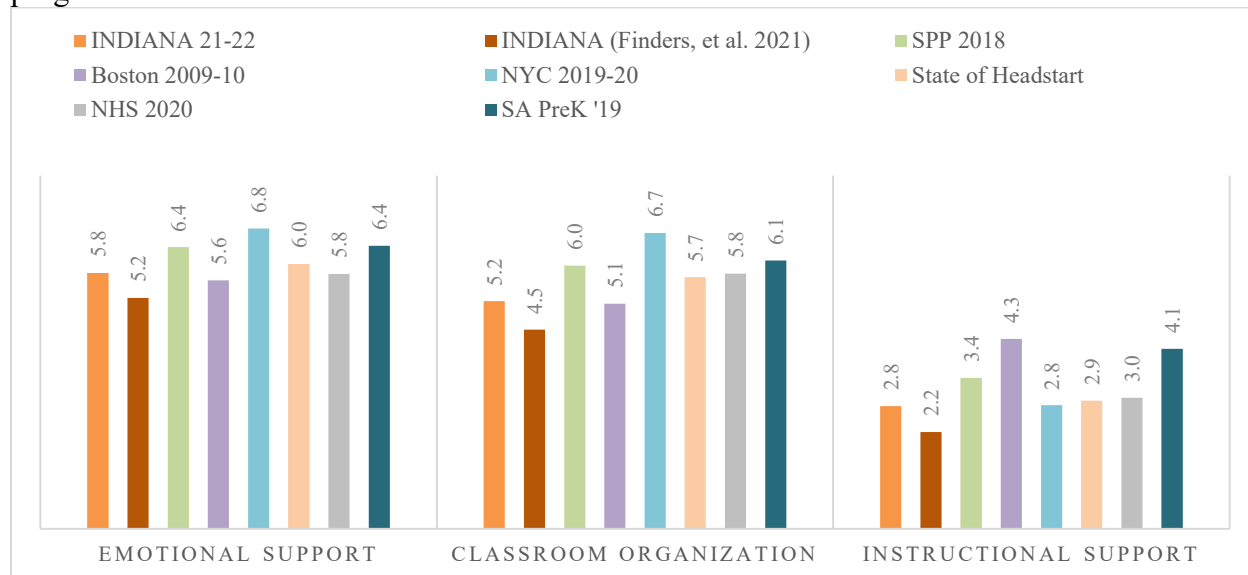
Figure 9. Average 2021 and 2022 CLASS Infant and Toddler ratings in comparison to selected state or program evaluations



Sources: Xue, et. al (2021) report Baby FACES 2018 results; Baby FACES results are reported in Xue, et., al (2022); Bandel, et. al (2014) report Baby FACES 2010-2012 results; Bichay-Awadalla & Bulotsky-Shearer (2022) for the toddler results reported; LA results are reported in LA Department of Education (louisinabelieves.com) and Bassok, et. al, 2021.

On the other hand, numerous state and city preschool programs have been studied using the CLASS Pre-K. Figure 10 displays the patterns of CLASS Pre-K scores for preschool classrooms in the study compared to those of other cities and states, including high-quality city-funded programs. It is noteworthy that the overall pattern of the Indiana sample is consistent with that of other studies, showing CLASS ES scores higher than CLASS CO scores, and significantly lower scores for CLASS IS. While the average CLASS ES scores in the Indiana sample are comparable to those of some other localities, such as Boston, they fall behind programs in New York City, San Antonio, Seattle (SPP), and other areas. For CLASS CO and CLASS IS, Indiana's scores are lower than most of the other programs depicted.

Figure 10. Average 2021-22 CLASS Pre-K scores in comparison to selected state or city program evaluations



Note: SPP is Seattle’s preschool program, reported in Nores, et. al (2019); NJ Abbot is New Jersey’s preschool program, reported in NIEER (2014), and SA Prek is the San Antonio PreK program in Decker-Woodrow, et. al (2019). State of Head Start is report in Barnett & Friedman-Krauss (2016). Finders, et. al (2021) reports other Indiana results. Boston results are reported in Weiland, et. al (2013). National Head Start scores are reported in Head Start, ECLKC (2020).

## 2. Teacher characteristics & self-reported quality

In addition to observational data, we also distributed a survey to teachers to collect information on their background and experiences, as well as teachers self-report on the quality of their teaching and learning practices and their assessment practices. We report on data collected from 119 teachers in 106 programs, of which the breakdown looked similar to the sample as a whole in terms of PTQ level, urbanicity, and program type. Findings from the survey are discussed below.<sup>11</sup>

### *Teacher socio-demographics, work experience, wages, and benefits*

Teacher qualifications and background experiences are reported in Table 5. The majority of teachers who responded to the survey were female (97.5%). They ranged in age from 17-77 years old, with a mean age of 41.2 years old. Most were white (88.1%), and had completed some college credits but no degree (27.4%), had an associate’s degree (22.2%), or had a bachelor’s degree (31.6%). Most speak only English (95.6%) fluently. Most (69.7%) reported they did not have any teaching certifications in the state of Indiana, although there were some differences by program type. Approximately 16.2% of teachers in center-based programs reported they had a teacher certification in the state, compared to 28.6% in home-licensed programs, 28.6% in

<sup>11</sup> Teacher surveys were distributed to all teachers in the NIEER sample, but do not include teachers in the Building Blocks program; 52.7% of the NIEER sample completed a survey. Teachers who completed the survey scored significantly higher on the Classroom Organization domain of the CLASS pre-K. The survey had slightly lower rates of completion than expected for teachers in 3-rated and center-based programs, and slightly higher for home-based teachers.

ministry programs, 28.6% in “other” types of programs, and 77.8% in LEAs (although just nine teachers in LEA programs returned the survey). There were also slight differences in teachers who responded who had a teacher certification in Indiana that favored 3- and 4-rated programs.

Years of teaching experience ranged from 0.33 to 46, with a mean of 13.3 years of teaching experience. Teachers also reported on the number of hours they work each week in their programs – responses ranged from 11-65 hours per week, with a mean of 40.48 hours.

Table 5. Teacher qualifications and years of experience

<i>Highest Level of Education</i>	N	Percent
Less than high school diploma	1	0.9%
High school diploma (or GED)	10	7.7%
Some college credits, but no degree	32	27.4%
Associate’s degree (two year)	26	22.2%
Bachelor’s degree	37	31.6%
Master’s degree	11	9.4%
Doctorate degree	1	0.9%
<i>Degree Area of Study</i>		
Early Childhood Education	34	50.7%
Education Other (e.g., Special Education, Elementary)	18	26.9%
Other (e.g., Arts, Psychology, Business)	15	22.4%
	Mean	SD
<i>Years of Experience</i>		
In education	13.27	10.1
In current position	8.6	8.7

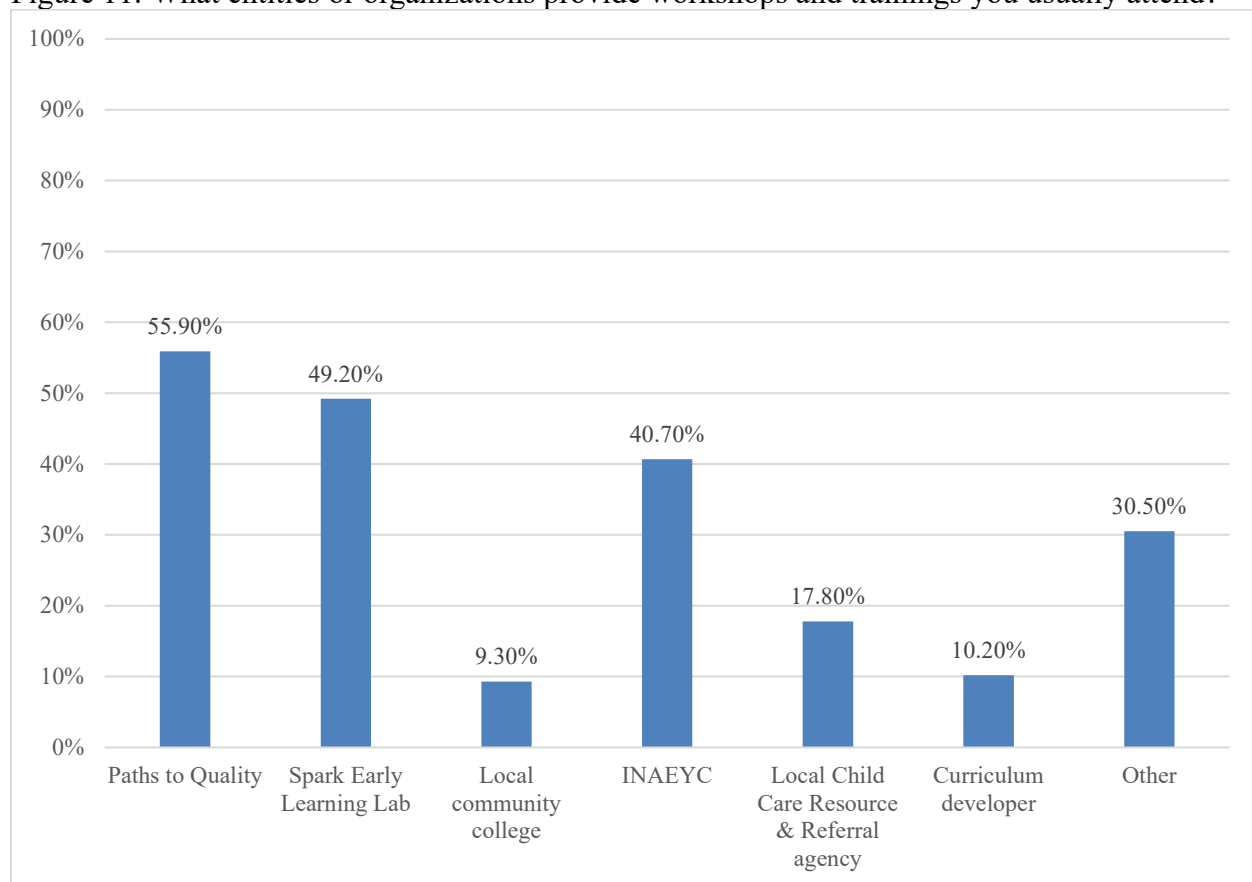
In relation to teacher salaries, 85% of the teachers in the sample reported annual total salaries of under \$50,000 a year. About 9% reported earning under \$10,000 a year, another 18% reported earning \$10,000-\$20,000 a year, 31% reported \$20,000-\$30,000 a year, 16% \$30,000-\$40,000, 9% \$40,000-\$50,000 and the remaining 17% over \$50,000. Of those reporting, 78% of teachers reported having paid vacations and holidays, 49% reported paid sick leave, 24% reported having full or partial health care, and 12% reported having a pension as part of their benefits. About 92% of the teachers reported an interest in continuing in this line of work in the forthcoming 3-5 years.

*Professional development opportunities*

To get a sense of teachers’ professional development opportunities, we asked them to report on whether or not they attend workshops, trainings, conferences, or classes offered by organizations/entities outside of their centers. The majority of teachers reported that they did this (81.9%). However, this differed by auspice. For example, 100% of teachers in FCC sites reported this, while this number was 66.7% for teachers in LEAs. There were also slight differences by PTQ level: 72.2% for non-rated, 70% for 1 & 2, 84.2% for 3-rated programs, and 92.3% for 4-rated programs. In terms of what entities or organizations provide these opportunities, teachers reported a mix of where they obtain this professional development,

highlighted in the below figure. Teachers could choose any options that apply. The most commonly selected option was Paths to Quality (55.9%), with Sparking Early Learning Lab (49.2%) and INAEYC (40.7%) following. Teachers could also select “Other” and write in where they obtained this professional development. In this space, most wrote about online opportunities such as Early Childhood Webinars and other online conferences. A few others mentioned I-LEAD and other district or corporate trainings. We also asked teachers, if yes, to respond with about how many hours per year they estimated they attended these trainings. The 89 teachers who responded provided a range of hours, from 0-140, with a mean of 23.4 hours. A small number of teachers mentioned they are completing degrees and presumably included these credit hours in their estimation of PD hours.

Figure 11. What entities or organizations provide workshops and trainings you usually attend?



### TSEEQ

TSEEQ average scores for the sample across all subscales are reported in Table 1. The patterns are consistent with the CLASS scores, with the highest scores recorded on the interaction and emotional climate subscale.<sup>12</sup> TSEEQ subscale scores range between 1 (minimal quality) and 5 (high quality) for all subscales, with the exception of the Physical Environment subscale, in which possible scores range from 1 to 4.5.<sup>13</sup> The average subscale scores for teachers in the

<sup>12</sup> This data is only available for about half the sample for which we got teacher survey responses.

<sup>13</sup> This scale only has a maximum of 4.5 by design given that various items in it are in a three-point Likert scale.

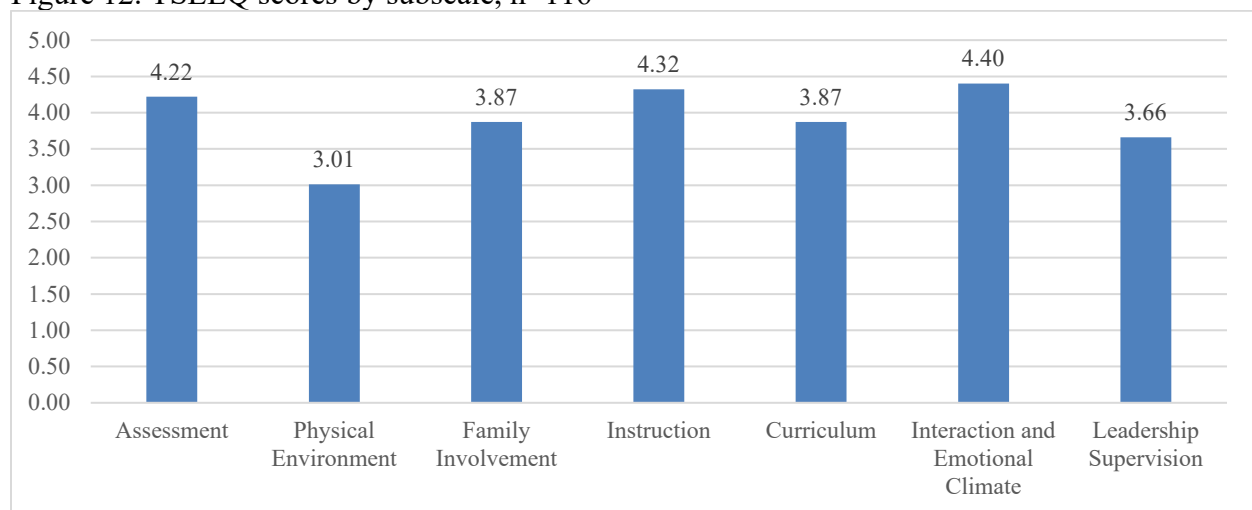
entire sample fall between 3.01 and 4.40, indicating a moderate to high level of quality in classrooms as reported by teachers.<sup>14</sup> The items used in each subscale indicate differing levels of engagement with high-quality practices across classrooms.<sup>15</sup>

On average, teachers reported lower levels of self-reported quality in relation to the physical environment (mean of 3.01), family involvement and the curriculum (means of 3.87), and leadership and related supports (mean of 3.66). In contrast, teachers report higher levels of quality in relation to assessment, instruction, and interactions/emotional climate.

Table 6. TSEEQ subscales and subscale means and ranges, n=116

TSEEQ Subscales	Mean	Standard Deviation	Minimum	Maximum
Assessment	4.22	0.64	2.00	5.00
Physical environment	3.01	0.33	2.00	3.50
Family involvement	3.87	0.69	2.00	5.00
Instruction	4.32	0.45	2.76	4.95
Curriculum	3.87	0.54	1.65	4.74
Interaction and emotional climate	4.40	0.43	3.00	5.00
Leadership supervision	3.66	0.60	1.46	4.79

Figure 12. TSEEQ scores by subscale, n=116



### *TSEEQ subscale scores*

Teachers scored the highest on the Interaction and Emotional Climate subscale, with a mean score of 4.40, or a score that indicates high quality. This subscale consists of 12 items in which teachers report on how often they engage in tasks such as comforting children when they are upset, spend extra time with new children who are transitioning into the classroom, or encourage

<sup>14</sup> We calculated subscale scores only for survey respondents who had completed at least half of the items within a subscale.

<sup>15</sup> We ran piece-wise correlations between the different age group domain scores for the CLASS and the domains for the TSEEQ. All correlations were low to moderate with no correlation above 0.35. This suggests that the TSEEQ and the CLASS assess different aspects of classroom quality.



children to respect each other's differences. Teachers also scored high on the Instruction subscale, with a mean score of 4.32. This subscale contains 18 items assessing how often teachers engage in practices like using play as an instructional strategy or planning and implementing small group activities. On the Assessment subscale, teachers responded to seven items about assessment practices, with an average subscale score of 4.22, or above average self-reported quality. Most teachers reported that they frequently or always engage in practices like assessing across multiple domains and documenting informal child assessment information. Teachers had a mean score of 3.87 on the Family Involvement subscale, which contains nine questions about how often teachers do things like hold special events at various times so multiple families can participate, and how often they encourage parents or other family members of different cultures/ethnicities to share cultural traditions. Teachers also reported a mean score of 3.87 on the Curriculum subscale, which contains 37 items focused on curriculum more generally and on different curricular domains, including literacy, math, social studies, and fine arts. On the subscales scored out of a possible 5.0, teachers scored lowest on Leadership Supervision, with a mean score of 3.66. This subscale contains 14 items and asks teachers to report how often they do things like reflect on their practice, have information shared with them from a supervisor/coach, and how often they attend workshops and trainings. Finally on the Physical Environment subscale, teachers scored 3.0 (out of a possible 4.5), which contains eight items about their physical environment, including whether or not they have materials that are in good condition and that living plants/animals are an everyday experience for children.

#### *TSEEQ subscale scores for selected center characteristics*

We also looked to see whether there were differences in self-reported quality as a function of various characteristics, including program type, PTQ rating, and county classification. These scores are reported in Table 7. As a whole, scores looked similar across these characteristics, with some differences noted. For example, teachers in low community poverty areas reported higher quality across all subscales of the TSEEQ than did teachers in areas with high community poverty –the score on the assessment subscale in low community poverty areas was 4.57, compared to 3.99 for teachers in high poverty areas. In general, teachers in LEA programs self-reported the highest quality, and higher subscale scores compared to teachers in licensed centers and ministry. Furthermore, teachers in 4-rated programs consistently reported higher quality compared to teachers in 1 & 2, 3, and non-rated programs. No differences emerged as a function of urbanicity.

Table 7. TSEEQ scores by subgroups, n=116

	Assessment	Physical Environment	Family Involvement	Instruction	Curriculum	Interaction & Emotional Climate	Leadership Supervision
<b>Facility Type</b>							
<b>Center-licensed</b>	4.35	3.00	3.92	4.36	3.95	4.43	3.60
<b>Home-licensed</b>	4.16	3.12	3.94	4.30	3.94	4.50	3.82
<b>LEA</b>	4.57	3.06	4.02	4.65	4.01	4.50	3.92
<b>Ministry</b>	3.99	2.88	3.68	4.17	3.65	4.23	3.50
<b>Other</b>	4.50	3.16	4.12	4.57	4.04	4.60	3.84
<b>Quality Level</b>							
<b>1 &amp; 2</b>	4.26	3.05	4.01	4.24	3.90	4.45	3.74
<b>3</b>	4.03	2.95	3.82	4.23	3.81	4.28	3.69
<b>4</b>	4.51	3.12	4.07	4.47	4.03	4.57	3.83
<b>Non-rated</b>	3.93	2.80	3.36	4.26	3.58	4.23	3.10
<b>County Classification</b>							
<b>Mid-sized</b>	4.10	2.91	3.77	4.31	3.84	4.38	3.46
<b>Rural</b>	4.20	3.07	3.87	4.29	3.82	4.36	3.69
<b>Urban</b>	4.29	2.99	3.92	4.35	3.92	4.45	3.72

Note: The following differences were statistically significant: Ministry-type facilities scored lower than Center-licensed, LEA, and Other type facilities on the Assessment scale; Ministry-type facilities scored lower than Home-licensed and Other type facilities on the Physical Environment scale; Ministry-type facilities scored lower than LEA and Other type facilities on the Instruction scale, Home-licensed facilities also scored lower than LEA facilities on the Instruction scale; Non-rated facilities scored lower than Level 4 facilities on the Assessment scale, and Level 3 facilities also scored lower than Level 4 facilities on the Assessment scale; Non-rated facilities scored lower than Level 1 & 2 and Level 4 facilities on the Physical Environment scale, and Level 3 facilities also scored lower than Level 4 facilities on the Physical Environment scale; Non-rated facilities scored lower than all other rated facilities on the Family Involvement scale; Non-rated facilities scored lower than Level 4 facilities on the Curriculum scale; both Level 3 and Non-rated facilities scored lower than Level 4 facilities on the Interaction & Emotional Climate scale; Non-rated facilities scored lower than all other rated facilities on the Leadership Supervision scale.

### DEC checklists

In addition to the CLASS and the TSEEQ, we used two performance checklists developed by the Early Childhood Technical Assistance Center (ECTA),<sup>16</sup> which are aligned with the DEC recommended practices.<sup>17</sup> The checklists were not developed for monitoring purposes, but provide the possibility of assessing variation in quality of processes across programs. We used two of these checklists. The Informed Clinical Reasoning (ICR) checklist focuses on “practices that are important for using informed clinical reasoning/opinion for evaluation and eligibility determination” for children. The Authentic Child Assessment (ACA) checklist includes “key characteristics of authentic assessment practices for observing child participation in everyday activities.” Both of these use a four-point Likert response system that assess frequency through the following responses: Seldom or never (0-25% of the time), Some of the time (25-50%), As often as I can (50-75%), and Most of the time (75%-100%). We first assessed whether the CLASS and checklists are capturing differing aspects of quality, as the checklists focused on the

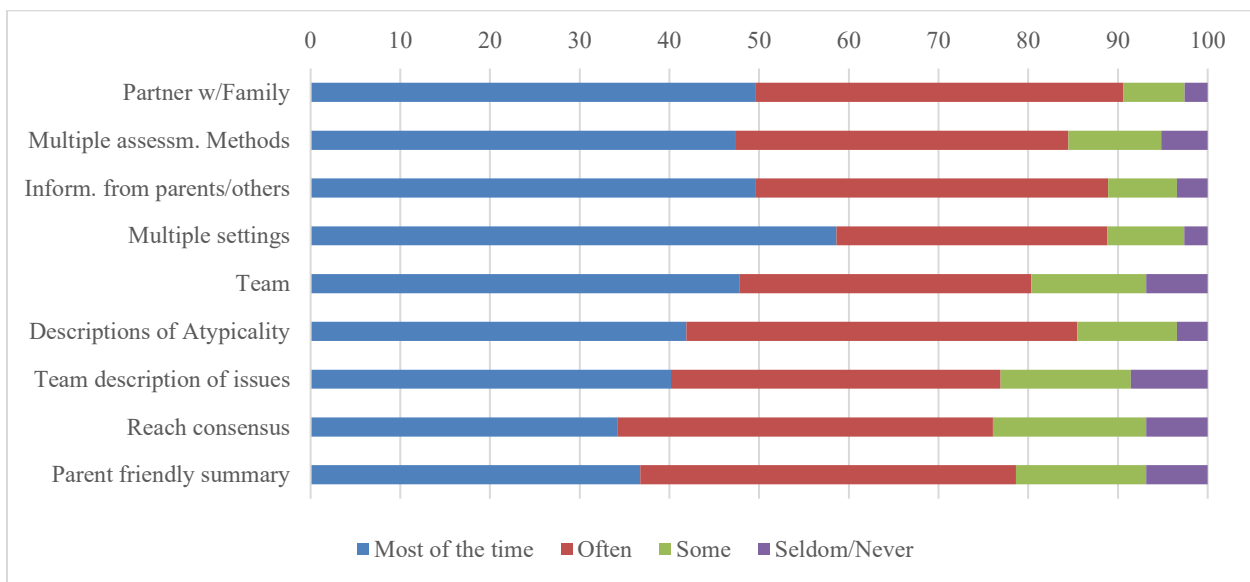
<sup>16</sup> <https://ectacenter.org/decrp/development.asp#analysis>

<sup>17</sup> Division for Early Childhood (2014). DEC recommended practices in early intervention/early childhood special education 2014. Retrieved from <http://www.dec-sped.org/recommendedpractices>

use of assessment and the process for identification, while the CLASS measures overall process quality.<sup>18</sup> We found this is, in fact, the case.<sup>19</sup>

In relation to the Informed Clinical Reasoning, most teachers described using the recommended practices often to most of the time (over 50% of the time) when assessing children’s needs. Figure 13 illustrates teachers’ overall responses in relation to these set of DEC practices. Around 20% of teachers reported not often engaging with a team in the process, in decisions related to children’s considerations for referrals and classification, and in communicating this to families in a parent-friendly way. Overall, between 40-50% of the teachers reported doing various activities related to informed clinical reasoning most of the time.

Figure 13. Informed Clinical Reasoning (ICR), frequency.

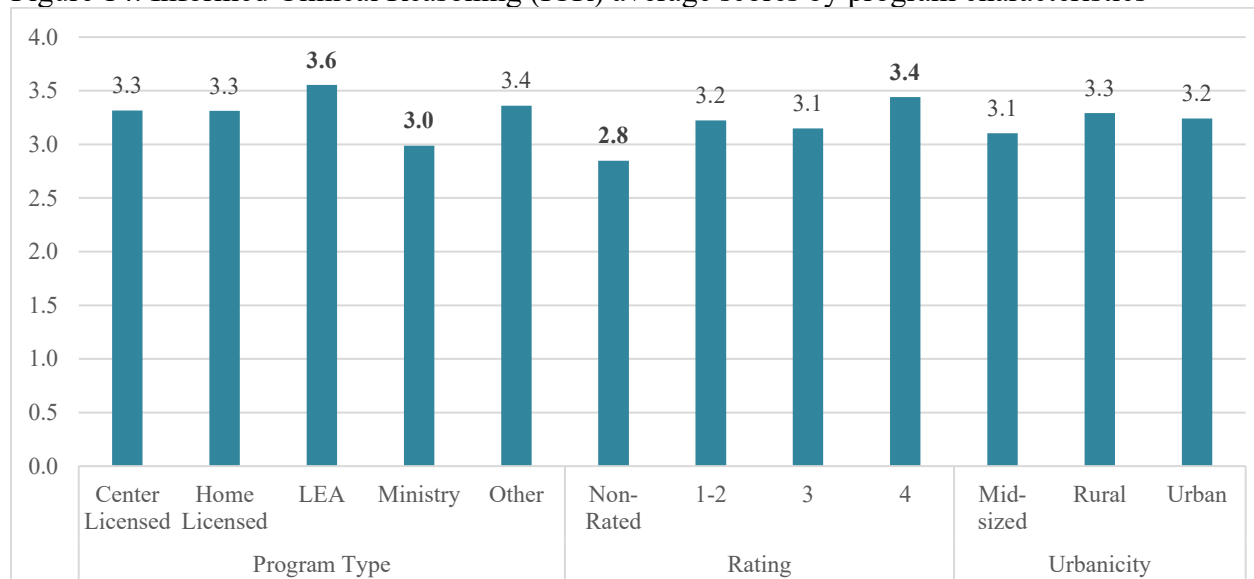


To compare programs, we calculated the average ICR score by rating the reported frequency from 1 to 4, with 1 representing the “Seldom/Never” category and with 4 representing the “Most of the time” category. Across program types, urbanicity, and rating characteristics, some small differences emerge. Teachers in LEA and 4-rated programs self-report higher levels of informed clinical reasoning, and those in Ministry programs report lower levels. The differences for the latter two are significant after controlling for differences in urbanicity, rating, and program type.

<sup>18</sup> We run pairwise correlations. Correlations between the two checklists were significant, while this was not the case between these and the CLASS infant, toddler, and pre-K scores.

<sup>19</sup> We ran piece-wise correlations between the different age group domain scores for the CLASS and the DEC checklists and found low correlations across the board. In addition, we ran piece-wise correlations between the TSEEQ and the DEC checklists to assess whether these capture different aspects of classroom quality. We found the TSEEQ domains and the total and averaged DEC scores for the two measures included were above 0.5.

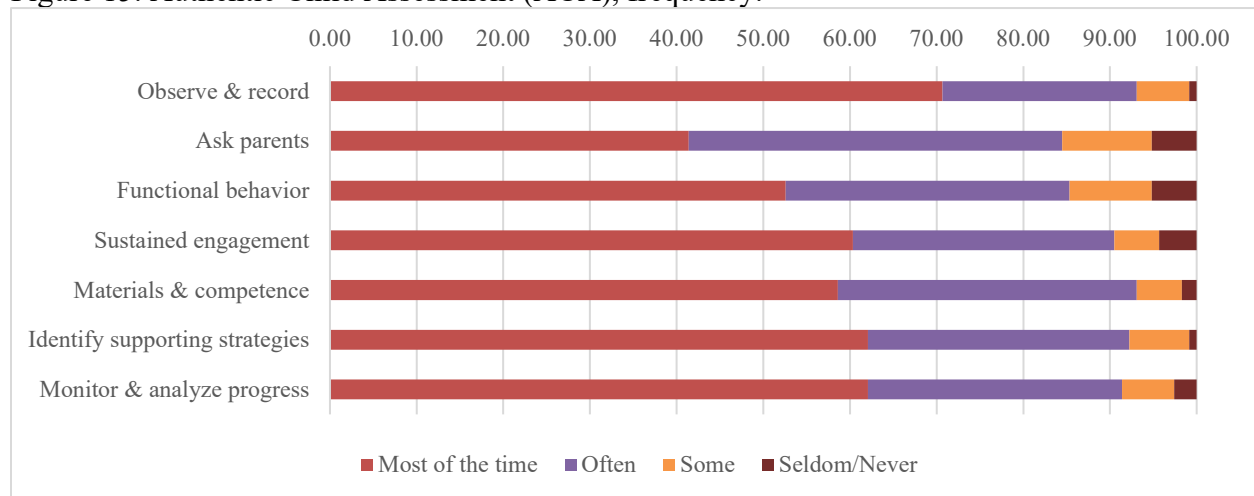
Figure 14. Informed Clinical Reasoning (ICR) average scores by program characteristics



Note: We estimated associations between the ICR score and program characteristics to verify which of the differences observed were significant. Those significantly different are bolded.

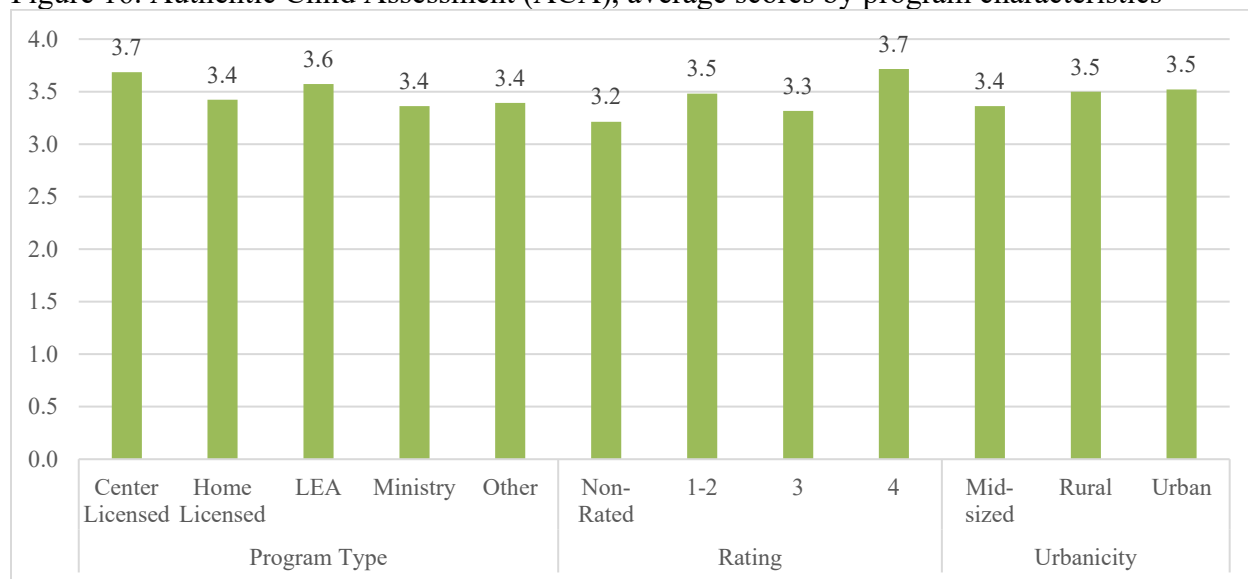
In terms of Authentic Child Assessment, teachers’ self-reports illustrated in Figure 15 below shows that about 60% of the teachers report engaging most of the time in the various activities related to observing children in naturalistic environments, recording their observations, understanding children’s behavior in relation to contexts and settings and on which aspects or materials they show sustained engagement, and consequently identifying next steps to support their development and monitoring.

Figure 15. Authentic Child Assessment (ACA), frequency.



Average ratings in the (ACA) by program characteristics are illustrated in Figure 16. Results show higher use of observation and assessment in center-based and LEA programs, as well as in programs rated 4, although these differences are not significant after controlling simultaneously for all program characteristics.

Figure 16. Authentic Child Assessment (ACA), average scores by program characteristics



Note: We estimated associations between the ICR score and program characteristics to verify which of the differences observed were significant. No significant differences emerged.

### 3. Children’s development in Indiana’s ECE programs

#### *Infants and toddlers*

This evaluation measured infant and toddler developmental gains in the areas of vocabulary (receptive and expressive), cognitive development, socio-emotional development, and were measured using the above-mentioned Bayley Scales of Infant Development IV. Infant and toddler scores for the 2021-2022 gains for the overall sample, and for selected subgroups of interest are shown below (reported in detail in Appendix C). Table 8 reports average standardized scores at baseline and in the follow-up, as well as gains. Gains are illustrated in Figures 17-20 for subgroups of interest. Gains allow interpreting the developmental progress of the children in the sample in relation to the typical average development of children of the same age. These measures are standardized at a mean (average) score of 100 and with a standard deviation of 15. Standard scores under 100 points signify developmental levels below average for children this age. Standard positive gains depict gains for children above those typical for children of the same age, and negative standard gains do not depict losses but rather a slower progression than average relative to children their age.

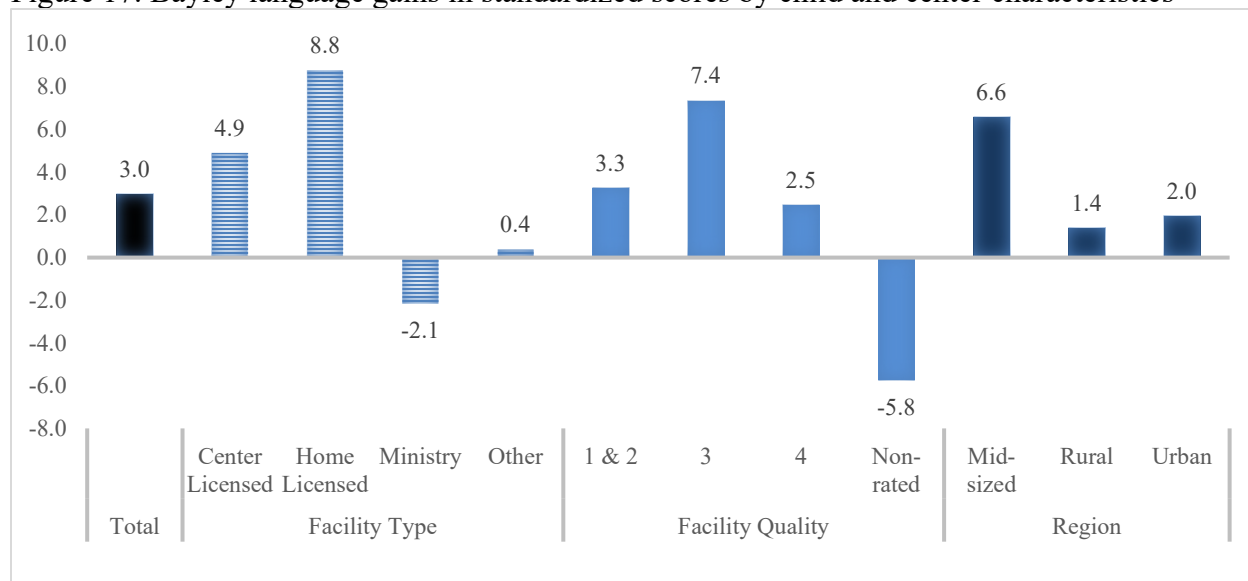
Table 8. Bayley gains in standardized score by child and center characteristics

	2021		2022		Gain*	
	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.
Language	89.6	18.3	91.8	16.0	3.0	18.6
Cognitive	93.9	16.8	96.3	13.8	2.5	17.0
Socio-Emotional	95.0	21.6	95.9	20.3	3.4	26.9
Socialization	100.1	16.5	99.5	12.4	0.9	16.8

Note: Gains are reported for those for which there is a pre-test and subsequently for those for which there is a post-test in the sample. Therefore, the mean will not be equivalent to the simple subtraction of the pre- and post-test mean. n=170-190 children for which there is pre- and post-test data, depending on the measure.<sup>a</sup> Statistically significant differences tests show that 2022 Cognitive and Language scores were both higher than 2021 scores (one-sided p-values).

Gains in language development for infant and toddlers are reported in Figure 17. Children in the sample average a baseline language standardized score of 89.6 (with a standard deviation of 18.3 and scores ranging between 45 and 130). That is, children in the study scored under the norm relative to children their age. Children in higher rated programs had slightly higher language scores at baseline, and baseline scores were highest for non-rated programs and in mid-sized communities (see report 1). Over the study period, children in the study gained in relation to typical peers their age, with mean standardized gains of three points in language. Gains were largest in home licensed (which had on average lower baseline scores), PTQ rated 3 (with low baseline scores as well), and mid-sized communities (which started with higher child scores on average and where children grew the most in terms of language).

Figure 17. Bayley language gains in standardized scores by child and center characteristics

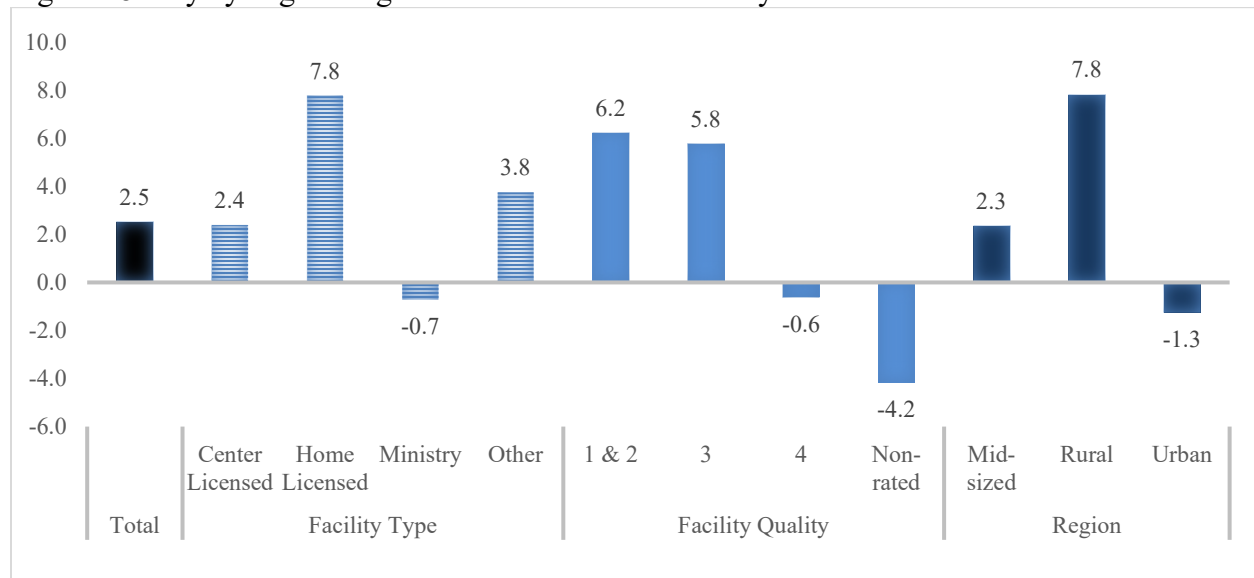


Note: We ran a test of equality of means in gains across groups using *MANOVA* analyses for each outcome. We find the following differences in mean gains - across urban classification groups, children from Ministry facilities gained less than those from Center-Licensed and Home-Licensed facilities. Children from Non-rated facilities gained less than those from rated facilities.

Standardized scores for cognitive development are shown in Figure 18. Infants and toddlers in the study scored on average under the levels expected due to maturation at both pre- and post-test. Children in the sample average a baseline cognitive standardized score of 93.9 (with a standard deviation of 16.8, and scores ranging between 55 and 130). Children in non-rated centers and centers rated at 4 evidenced average higher baseline levels in cognitive development

in relation to peers their age and in the study. Average cognitive standardized scores for children were also higher in programs in mid-sized communities and much lower in rural communities. Over the study period, children in the study gained in relation to typical peers their age, with mean standardized gains of 2.5 points in cognition. Children in rural communities started with the lowest baseline scores and grew the most on average. Similarly, children in home-licensed care evidenced the lowest baseline scores and grew the most. This is also the case for cognitive gains for children in centers with a PTQ rating of 3 or under.

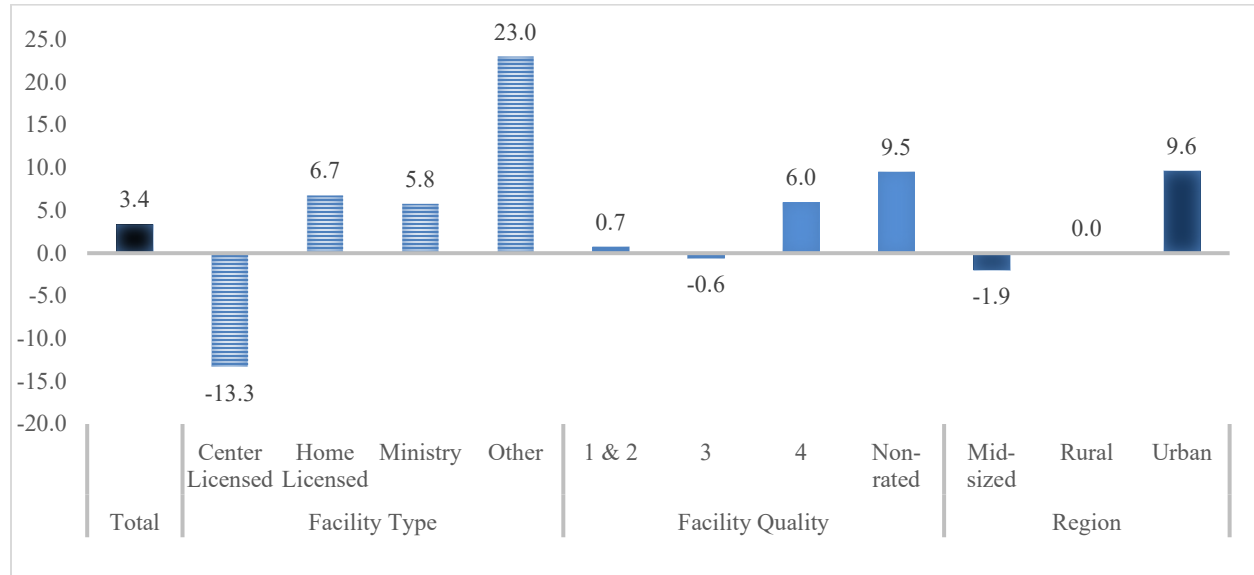
Figure 18. Bayley cognitive gains in standardized scores by child and center characteristics



Note: We ran a test of equality of means in gains across groups using *MANOVA* analyses for each outcome. We find the following means are statistically different - urban classification groups, children from Non-rated facilities gained less than those that from Level 1 & 2 and 3 facilities, children from rural areas gained more than those from mid-sized and urban areas.

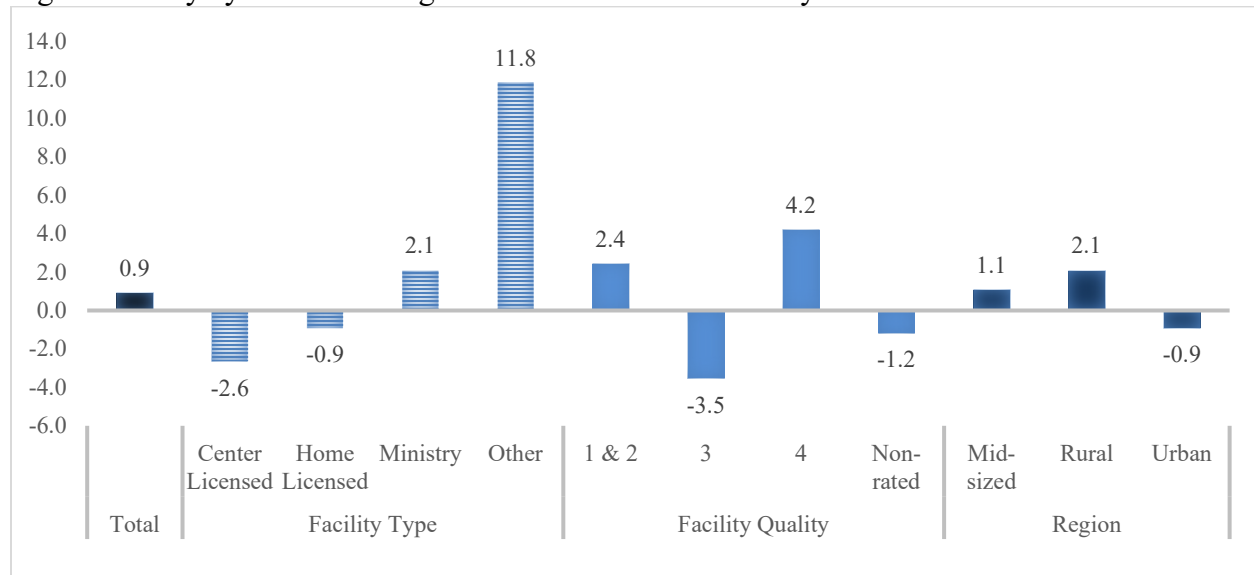
Figure 19 reports average standardized scores for teacher-reported socio-emotional development and Figure 20 reports these for socialization. Infants and toddlers in the study scored slightly under or at average levels relative to their peers at baseline. Between pre- and post-test children in the sample improved slightly in terms of socio-emotional development relative to what would be expected for their age, but mostly stayed on par with peers in socialization. Scores varied slightly across quality levels, with higher growth in children in centers rated 4 or unrated in socio-emotional development, and in centers rated 4 in socialization. It is important to note that these scores are based on teacher reports where there was a low response rate and a high amount of variance (high standard deviations) by teachers at post-test.

Figure 19. Bayley socio-emotional gains in standardized scores by child and center characteristics



Note: We ran a test of equality of means in gains across groups using *MANOVA* analyses for each outcome. We find the following means are different - children from Center-licensed facilities gained less than those from other types of facilities.

Figure 20. Bayley socialization gains in standardized scores by child and center characteristics



Note: We ran a test of equality of means in gains across groups using *MANOVA* analyses for each outcome. No statistical difference was found on means across urban classification groups.

### ***Multivariate analyses for infant and toddler child gains***

We also assessed the degree to which classroom and center characteristics are related to infant and toddler gains, controlling for child socio-demographic characteristics. We conducted multivariate estimates to examine the association between children’s gains and types or location of programs. Results showed few differences in children’s development by type of program (i.e.,



LEA, center-based, home, other), by urbanicity or by PTQ rating. However, the trends were not systematic across all child outcomes. Exploratory analyses suggest a negative association between “ministry” programs and children’s language gains, and a positive and significant association for programs with PTQ ratings of 3 and 4 and children’s language gains. The positive association was also present for programs with a rating of 3 in terms of cognitive gains.

### *Preschool-age children*

This evaluation measured gains in children’s receptive vocabulary (using the Peabody Picture Vocabulary Test), literacy (using the Woodcock-Johnson Tests of Achievement Letter-Word subtest), and math (using the Woodcock-Johnson Tests of Achievement Applied Problems subtest). Moreover, it evaluated executive functioning (EF) using two measures: the Dimensional Change Card Sort Game (DCCS) and the Peg Tapping task (PT). Socio-emotional development was measured with the ASEBA teacher reported form (C-TRF) as described above.

Developmental gains for the 2021-2022 follow-ups across selected subgroups of interest are shown below and reported in detail in Appendix D. Table 9 reports baseline and follow-up standardized scores for all measures, in addition to pre-post gains. Vocabulary, literacy, and math measures are standardized (at a mean score of 100 and with a standard deviation of 15). Gains (or losses) should be interpreted as changes relative to peers their age.<sup>20</sup> Children in the sample showed average baseline scores above children in the fall 2019 FACES study of Head Start, where children scored on average at 81.4 (with a SD of 14.8 and a range of scores of 40 through 132).<sup>21</sup> In literacy, children also had baseline scores above the 2019 FACES Head Start study sample (which scored at an average 86.4 with a standard deviation of 13.4 and a range between 52 and 148).<sup>22</sup> For math, average baseline scores were also above the 2019 FACES Head Start study sample (which scored at an average 80.3 with a standard deviation of 16.1 and a range between 41 and 126).<sup>23</sup>

Table 9 reports the PPVT IV (vocabulary) and Woodcock-Johnson (literacy and math) gains made by children in the sample in comparison to the average gains made by children their age.<sup>24</sup> Since measures are standardized, positive gains signify growth larger than that of typical children of the same age and negative gains should be interpreted as slower than average progress. As reported earlier, children started off with scores above average for their age (and above those in the Head Start study reported earlier) and they gained in relation to the growth of typical peers their age in vocabulary. In literacy, standard gains were negative, which implies smaller gains that were not at par with those of average peers their age. For math, gains were slightly negative,

---

<sup>20</sup> As the instruments were standardized with normative samples several years back, the standardized scores and gains can feasibly help understand the sample in the study relative to pre-pandemic conditions.

<sup>21</sup> See Table B.4. on page 80 in Kopack Klein, A., Aikens, N., Li, A., Bernstein, S., Reid, N., Dang, M., Blesson, E., Rakibullah, S., Scott, M., Cannon, J., Harrington, J., Larson, A., Malone, L., Tarullo, L. *Descriptive Data on Head Start Children and Families from FACES 2019: Fall 2019 Data Tables and Study Design. OPRE Report #2021-77.* Washington, DC: U.S. Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research, and Evaluation. <https://www.acf.hhs.gov/opre/report/descriptive-data-head-start-children-and-families-faces-2019-fall-2019-data-tables-and>

<sup>22</sup> Ibid. See Table B.8. on page 84.

<sup>23</sup> Ibid. See Table B.11. on page 87.

<sup>24</sup> This is referred to as standard score gains.

but close to zero, which means children’s gains were on par with typical peers their age. For the literacy and math measures, children in this sample were performing under the norming sample at pre-test, and with a slower progress than typical for children of similar age; they were found to be further behind by the spring assessment.

This table also reports pre-test, post-test, and gains in the DCCS and Peg Tapping (executive function) and the C-TRF measure (socio-emotional). Children in the sample averaged a pre-test scaled score on the DCCS of 1.7 (with a standard deviation of 0.8 and a range between 0 and 15). As referenced, the Learning-Related Cognitive Self-Regulation School Readiness Measures for Preschool Children Study (aka the Self-Regulation Measurement Study; [Meador, et al., 2013]) reports average DCCS scores of 1.42 at 51–53 months of age and 1.62 at 57–59 months. This is an average difference of 0.20 between these two ages. Children in this IN sample gained in the DCCS and PT similarly to the referenced study, with overall DCCS scaled gains being 0.2. The self-regulation study also shows PT average scores of 6.02 at 51–53 months and 8.80 at 57–59 months, with a difference of 2.78. For Peg Tapping, children in the sample averaged at pretest a score of 7.5 (with a standard deviation of 6.3 and a range between -1 and 16) and gained on average 1.9 points. These are lower gains in relation to those in the referenced study. Other studies in Seattle and Boston have found higher gains for children (Nores, et al., 2018; Weiland & Yoshikawa, 2013; Weiland, et al., 2013).

As pertains to children’s socio-emotional development, the C-TRF measure captures the presence of issues and changes should therefore be interpreted inversely. That is, increases in the scores imply a higher incidence of issues, while reductions in the scores translate into a reduction in the incidence of problems in the classroom. The scores reported reflect how a child’s score on each scale compare with the scores of the normative sample of peers. The incidence of socio-emotional problems decreased between the pre and post-test with most improvements observed in externalizing behaviors. The measure is reported by teachers and there was a large reduction in the number of teacher reports (with only 32% of the sample having post-test scores). Therefore, these results should be interpreted with caution.

Table 9. Average 2021 and 2022 scores and gains for children ages 3-5

	Pre-test 2021		Post-test 2022		Gains*	
	Mean	SD	Mean	SD	Mean	SD
<b>PPVT</b>	102.4	16.3	118.7	21.9	17.0	15.5
<b>WJ-AP Literacy</b>	93.1	12.7	91.0	12.6	-1.8	6.5
<b>WJ-AP Math</b>	94.9	14.6	93.7	8.0	-0.4	10.8
<b>DCCS Raw*</b>	13.7	6.1	15.0	5.7	1.4	4.0
<b>DCCS Scaled</b>	1.6	0.6	1.8	0.6	0.2	0.6
<b>Peg Tapping</b>	7.9	6.3	9.2	5.8	1.9	4.6
<b>Socio-Emotional Problems</b>	49.0	11.0	49.4	11.6	-1.0	8.1
<b>Internalizing Problems</b>	47.2	10.3	49.0	10.8	-0.1	8.7
<b>Externalizing Problems</b>	50.8	10.3	50.4	10.5	-1.2	6.9

Note: n=499-506, depending on the measure. The gains reported are for the analytical sample for which there is a pre and post-test. For socio-emotional, this was dependent on teacher response, which was high at baseline but low at post-test. Gains are therefore only available for 32% of the children. We excluded extreme values under or above 1.5 SDs. \*Interquartile range within each measure. DCCS Raw scores exclude Building Blocks score as these were shared as scaled; however, patterns in raw and scaled scores resemble each other.

Figure 21 illustrates the gains reported above in standard scores for receptive vocabulary, language, and math, and in raw and scaled scores for executive functions. Overall, children in the sample average gained in receptive vocabulary in relation to peers their age 17 standard score points (which translates into a 1.10 standard deviation gains). They lost 1.8 standard score points in literacy (which translates into a 0.28 standard deviation loss) and lost 0.4 standard score points in math (which is equivalent to -0.04 standard deviations). The second panel of Figure 21 illustrates the DCCS raw and PT raw gains. As reported above, the growth in these appears to be at par or slightly under those typical of peers this age.

Figure 21. Standard score gains in Vocabulary, Literacy, Math, and Executive Function

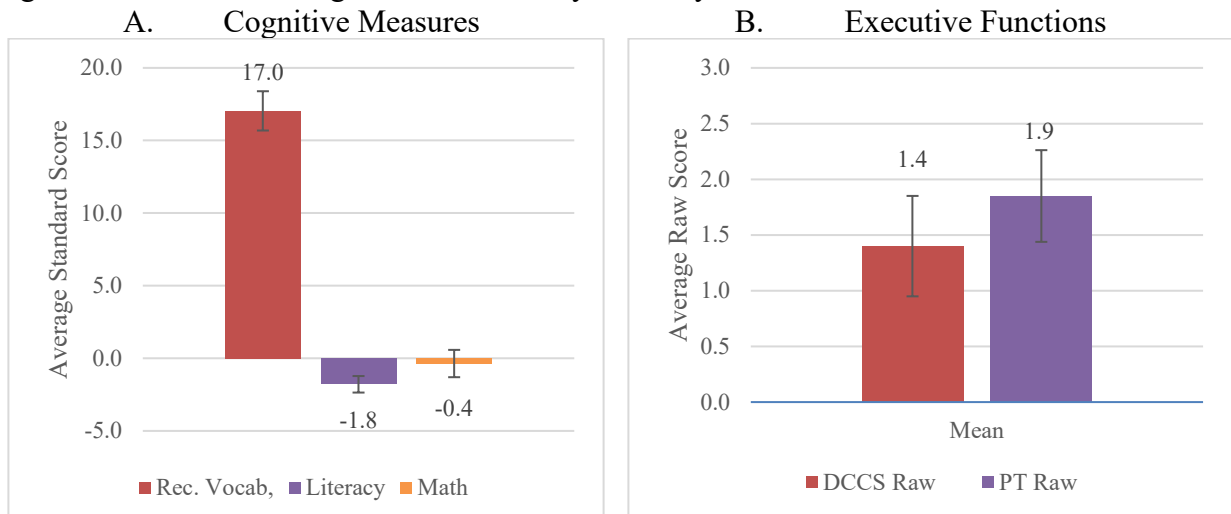
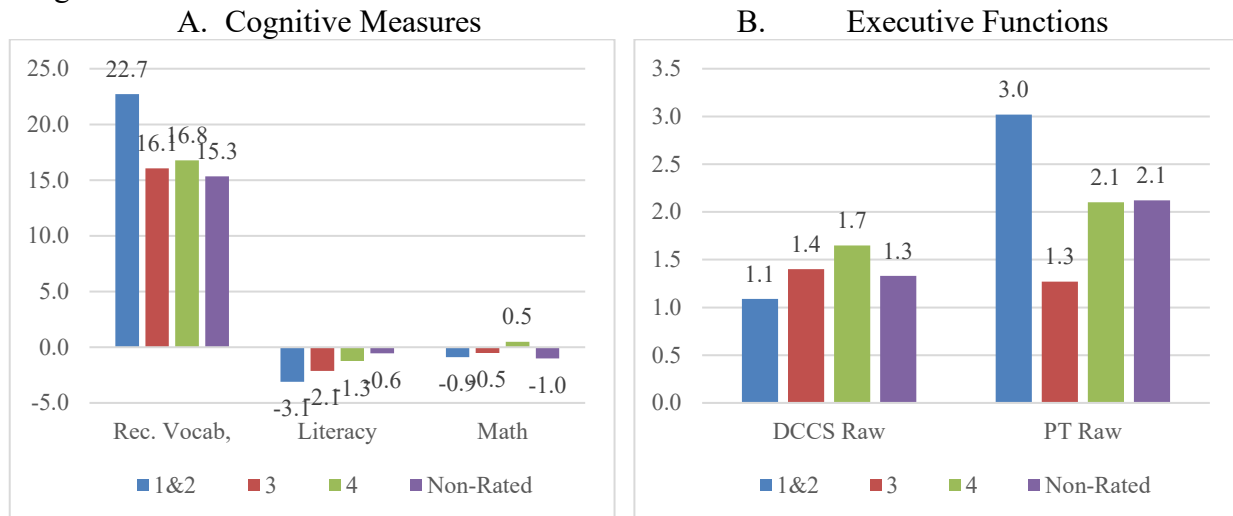


Figure 22 replicates these results by PTQ rating. On average, children in centers with the lowest ratings (1 & 2) showed significantly higher gains in vocabulary. In contrast, children in these centers also had a larger literacy loss relative to typical peers their age. In relation to executive functions, all types of centers showed similar growth in these, with some differences across the two measures. Overall patterns of large vocabulary gains, negative literacy ones, and gains in math and executive functions in line with similarly aged peers are present across all PTQ ratings. Only the differences across groups for PPVT and Peg Tapping were statistically significant.

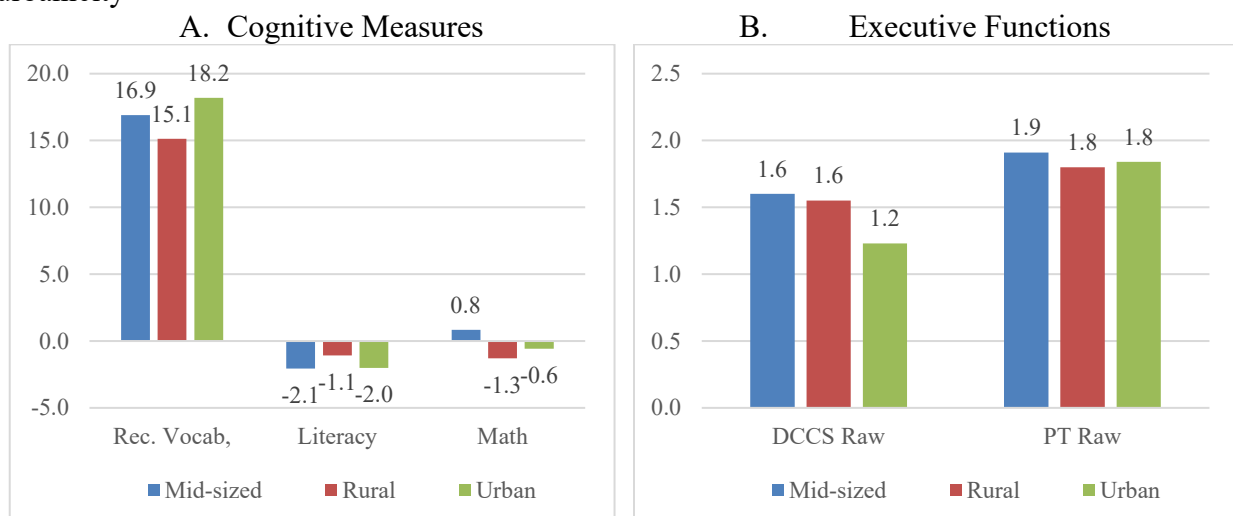
Figure 22. Standard score gains in Vocabulary, Literacy, Math, and Executive Function by PTQ rating



Note: We ran a test of equality of means in gains across groups using *MANOVA* analyses for each outcome. We conclude the following means are likely different across groups at a 95% confidence level: children from Level 1 & 2 facilities gained more than those from other facilities on PPVT; children from Level 1 & 2 gained more than those from Level 3 facilities on Peg Tapping.

Figure 23 reports changes in the different measures by urbanicity. Children in centers in urban areas seem to have shown slightly higher growth in vocabulary between pre- and post-test (and children in urban centers also averaged slightly lower baseline scores to start with). Children in urban centers also showed lower gains in the DCCS measure of executive functions, with no other differences emerging by urbanicity. Patterns of high growth in vocabulary and executive functions relative to peers of similar age, and similar or slightly lower growth in literacy and math are similar for children regardless of urbanicity. The overall trends reported generally and by PTQ ratings are also evident by urbanicity.

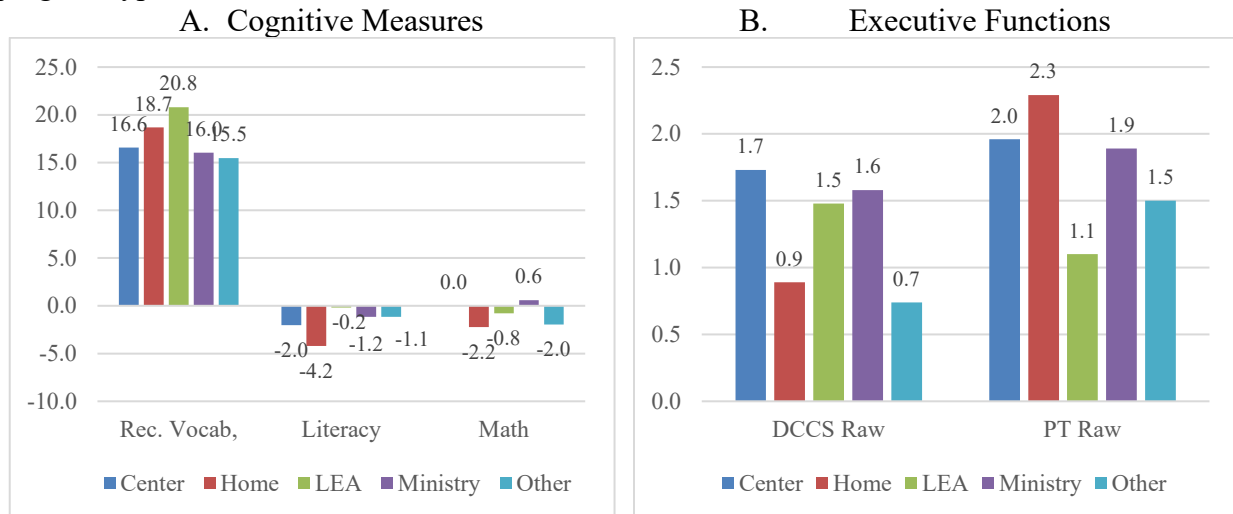
Figure 23. Standard score gains in Vocabulary, Literacy, Math, and Executive Function by urbanicity



Note: We ran a test of equality of means in gains across groups using *MANOVA* analyses for each outcome. We conclude means were not likely different across groups at a 95% confidence level.

In relation to program type, while overall patterns resemble those above, some differences emerged. Gains in vocabulary were the highest for LEA programs. In addition, children in these programs showed the growth in literacy quite on par with typical gains of peers of similar age, unlike the rest of the programs where children showed slower growth. In addition, children in center-based programs show evidence of slower growth in math relative to peers. In contrast, children in these programs and Ministry programs show growth in executive functions consistent across the two measures.

Figure 24. Standard score gains in Vocabulary, Literacy, Math, and Executive Function by program type



Note: We ran a test of equality of means in gains across groups using *MANOVA* analyses for each outcome. We conclude means were different across following groups at a 95% confidence level - children from Home-licensed facilities gained less than those from all other type of facilities on Literacy.

As mentioned earlier changes in children’s socio-emotional development, while small in relation to standards, seem to indicate an improvement in behavior issues (due to the reduction in scores). Improvements were larger in externalizing behaviors in relation to the norming sample

(Appendix C). Improvements in internalizing and externalizing behaviors appear to be larger in programs rated 3 and 4, and in rural programs.<sup>25</sup>

Figure 25. C-TRF changes in internalizing and externalizing behavior.



Note: We ran a test of equality of means in gains across groups using *MANOVA* analyses for each outcome. We conclude means were different across the following groups at a 95% confidence level - children from mid-sized region gained more than children from other regions on externalizing behavior scale.

### Multivariate analyses for preschool child gains

In addition to looking at overall gains by subgroups of children, we also assessed the degree to which classroom and center characteristics are related to children gains, controlling for child socio-demographic characteristics. We therefore conducted multivariate estimates to examine the significance in child gains of being enrolled in different program types and settings. Results showed few differences in pre-K children’s development by type of program (i.e., LEA, center-based, home, others), by urbanicity or by PTQ rating. However, the trends were not systematic across all child outcomes. The associations of this type that emerged significant (and in fact, were negative) were for receptive vocabulary for children in LEA programs, as well as rural and urban programs.<sup>26</sup> These children therefore showed smaller growth during this period.<sup>27</sup> PTQ rated 3 programs also evidenced lower receptive vocabulary score gains and rural programs evidenced lower EF gains in the DCCS, while PTQ Rated programs 3 and 4 evidenced lower EF gains in the PT measure. Surprisingly, ratings on CLASS Emotional and Instructional Support in

<sup>25</sup> Response rates by teachers on socio-emotional child outcomes was lower at post-test. We tested differences in children for which we had and did not have data at post-test. There were no statistically significant differences in age and gender. However, the group with missing data was more likely to be non-white.

<sup>26</sup> Children in LEA programs had higher scores at baseline.

<sup>27</sup> CLASS Emotional Support scores were also positively associated with the DCCS executive function gains in the raw measure, although these do not include Building Blocks sites. Females seem to systematically perform better in literacy, and older children in receptive vocabulary (while statistically significantly performing less well than their younger peers in math). White children evidence statistically significant higher gains in receptive vocabulary and lower gains in math.

preschool classrooms were negatively related with receptive vocabulary gains, while ratings in CLASS Classroom Organization were positively associated with receptive vocabulary and literacy gains. With the few exceptions mentioned, there were no systematic pattern of differences in gains across these subgroups of interest of programs in literacy, math, and EF measures.

#### 4. Parent surveys

We also collected data from parents/caregivers on their perceptions of what constitutes a quality ECE program. We asked parents to report on their preferences regarding child care/preschool programs, the resources, and activities they have available in the home, and their perceptions regarding teacher supports to children. We report on data collected from the parents of 324 children in 134 different programs, including 97 infant/toddler parents and 227 pre-K parents.<sup>28</sup>

Surveys were mostly filled out by mothers (89.1%) with a small amount filled by fathers, foster parents, grandparents, aunts or uncles, or others. Focal children were roughly evenly split between males and females; 79.6% are white, 9% Black, 4.6% multi-racial, 2.8% Hispanic/Latino, 0.6% Asian, and 3.4% other. Responding families were distributed across program types: 34.6% had children in ministry programs, 29.6% had children in center-licensed programs, 22.2% in FCCs, 9.6% in LEAs, and 4% in programs categorized as other.<sup>29</sup> In terms of the regional break down, 39.5% of parents were in urban areas, 32.1% rural, and 28.4% mid-sized.

About 40% of parents reported incomes greater than \$100,000 per year. Another 29.3% of the sample reported household incomes in the \$50,000 - \$99,999 range, and 22.7% reported household incomes under \$40,000 per year (about half of which reported household incomes between \$30,000 and \$40,000). To further understand the sample in terms of available resources, we asked parents whether or not they received five forms of governmental assistance; 20% reported receiving Women, Infants and Children (WIC), 35% reported receiving Medicaid, and 37% reported benefitting from a Child Care Subsidy (e.g., CCDF, On My Way Pre-K, or Build Learn Grow Scholarship).

The majority of the parents (66.7%) reported they were between ages 30-39. Most had either completed a bachelor's (32.3%) or master's (26.3%) degree as their highest level of education. In contrast, 10.2% had a high school diploma or less. Findings from the survey are discussed below. Refer to tables D.1 – D.4 in the appendix for a detailed breakdown.

---

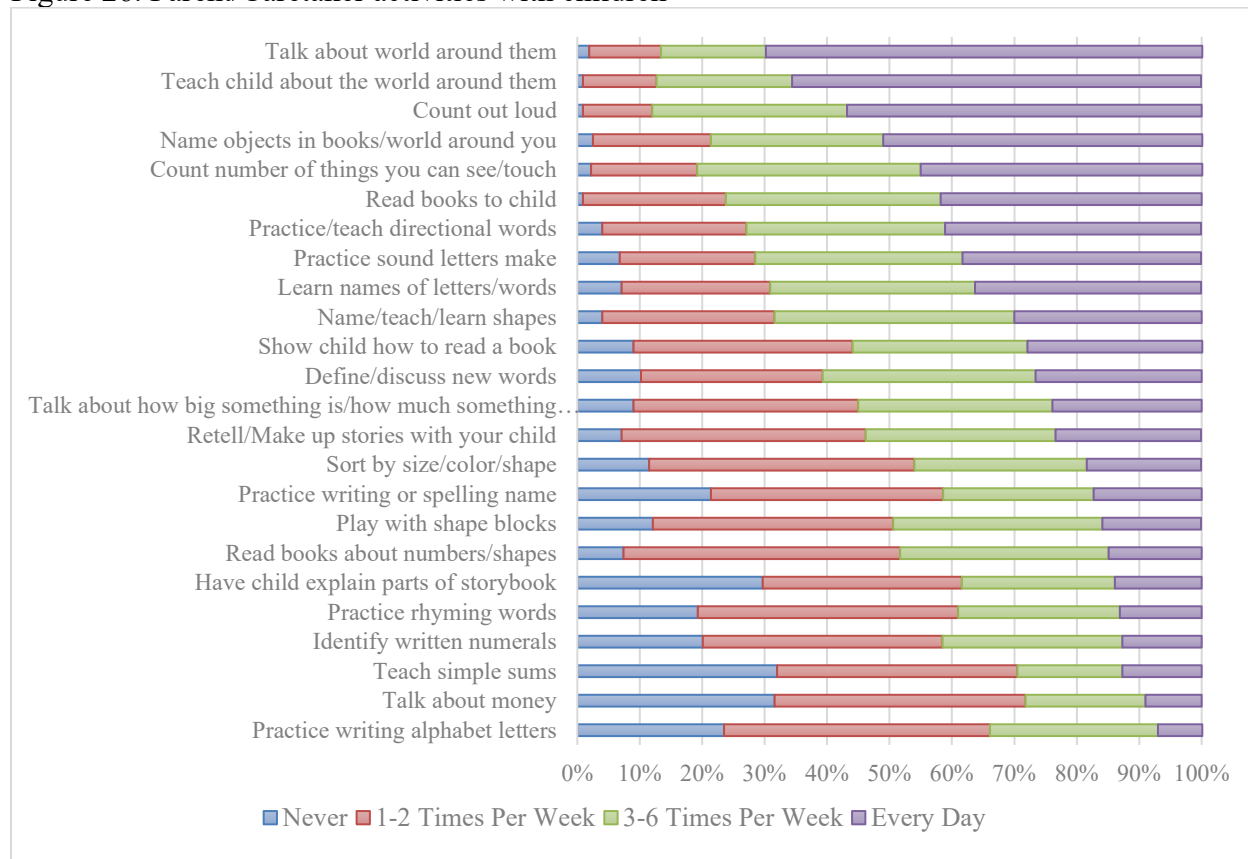
<sup>28</sup> Parents of children in 4C programs did not complete parent surveys; 46.6% of infant/toddler parents in the NIEER sample and 46.14% of pre-K parents in the NIEER sample completed the parent survey.

<sup>29</sup> Programs noted as “other” are those that are not classified as eligible to be included in the Paths to Quality QRIS (mostly public and private school-based programs that are legally license-exempt).

*Resources and activities in the home*

We asked parents/caregivers to report on the frequency with which they engaged in a number of different activities with children. Parents could select response options ranging from never to every day. Parents most commonly reported that they talk to children about the world around them every day (69.9%) and teach children about the world around them every day (65.5%). Conversely, less frequent every day behaviors were teaching their children simple sums (32.0%), talking about money (31.6%), or having children explain parts of a storybook (29.7%). As expected, there were some age-related differences in these reports. For example, both parents of infants/toddlers and pre-K-aged children reported they most frequently (i.e., every day) talk to children about the world around them (about 70%); but rates of defining/discussing new words with children differed somewhat as a function of age: 6% of pre-K caregivers reported they never do this, while 18.6% of infant/toddler parents report they never do this. And while 46% of pre-K age parents reported that they name objects in books/the world around their children every day, 63% of infant/toddler parents doing so every day.

Figure 26. Parent/Caretaker activities with children



Note: The survey asked: “In a typical week, how often do you do the following activities with your child?”

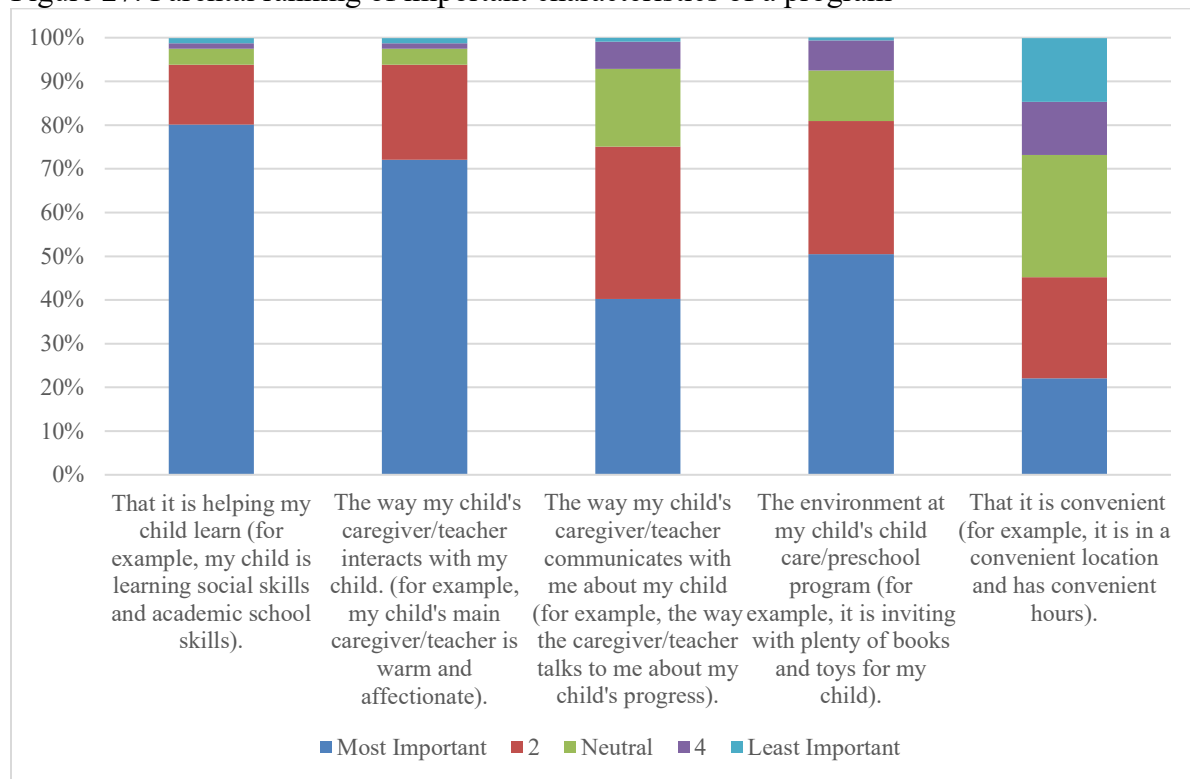
*Parental preferences on child care or preschool program*

Next, we asked parents to reflect on what was most (and least) important to them about their children’s child care/preschool program, ranking items as most important (1) to least important



(5). Parents most frequently reported that the most important characteristic to them about their child’s program was that it was helping their child learn (80% of parents ranked this as most important). The next most important characteristic reported by parents was that the way in which caregivers/teachers interact with children, including being frequently warm/affectionate, with 72% of parents ranking this as most important. In contrast, only about a fourth of parents ranked cost and convenience as most important.

Figure 27. Parental ranking of important characteristics of a program

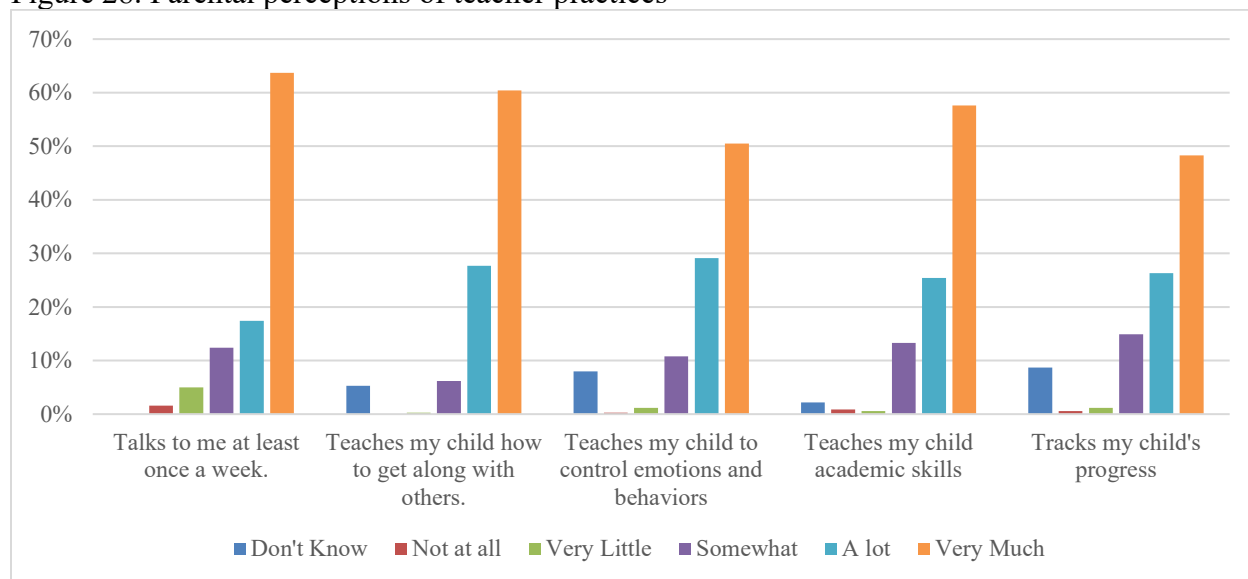


Note: The survey asked: “Rate what is MOST important about your child's preschool/child care program.”

### *Parental perceptions on teachers’ support to their children*

In terms of their confidence in the support teachers were providing to children, most parents/caregivers in the sample reported feeling their children were being supported in their classrooms. We asked parents/caregivers to respond to six statements on how their child’s teacher supported their child’s growth and development. For all six indicators, more than 50% of the sample indicated that they “very much” agreed with the statements. These responses somewhat ranged. For example, while under 50% of parents indicated that they “very much” agreed their child’s teacher tracked their child’s progress, about 60% of parents indicated they “very much” agreed with the statement that their child’s teacher met with them once a week and taught their child how to get along with others. Responses for the full sample can be seen in Figure 28 below.

Figure 28. Parental perceptions of teacher practices



Note: Parents were also asked if they felt their child’s teacher is fluent in the child’s primary home language; the majority agreed with this statement, which was expected due to the nature of the sample.

Some responses to these questions varied slightly by children’s age. For example, 60% of pre-K parents indicated that their child’s teacher talked to them at least one time per week at “very much;” 71% of parents of infants/toddlers felt this way. And while 63% of pre-K teachers indicated they very much agreed with the statement that their child’s provider “teaches my child academic skills,” 45% of infant/toddler parents felt this way.

There were also some differences in parental perceptions of support to children as a function of PTQ level. These are highlighted in table E.2. For example, between 64.5-70% of parents in 1-4 rated programs indicated they “very much” agreed that their child’s teacher talks to them at least one time per week, but this was only the case for 46.5% of parents in non-rated programs. In addition, while 57.6-68.4% of parents in 1-4 rated programs indicated they “very much” agreed that their provider teaches their child to get along well with others, this was the case for 41.9% of parents of children in non-rated programs. And 25.6% of parents in non-rated programs indicated they “very much” agreed that their child’s teacher was teaching their child to control emotions and behaviors, compared to 44% of parents in 1- and 2-rated programs, 60.1% of parents in 3-rated programs, and 51.1% of parents in 4-rated programs. Finally, differences emerged in how much parents perceived their children’s teachers were teaching their child academic skills; for example, 51.2% of parents in non-rated programs agreed “very much” with that statement, 42% of parents felt this way in 1- and 2-rated programs, 61.6% in 3-rated programs, and 63% in 4-rated programs.

*Parental perceptions on what constitutes a high quality early child care and education*

Finally, we asked parents to report on which components they believe make up quality in the ECE setting. Specifically, we asked parents: “When you think about what best represents quality care for an early care and education program, which of the following would be the most important to you personally? (Please select your top three).” Parents were given 13 options to

choose from, ranging from using a curriculum to communicating with parents. We aggregated results from 295 parents.<sup>30</sup>

The most frequently selected option was: “Providers are warm, caring, and provide empathy to students;” a total of 54.6% of respondents selected this option. The least important components for families were “Education programs follow national health and safety best practices” (7.8% of the sample) and “Provides careful supervision adjusted for different ages and abilities” (8.1% of the sample).

Table 10. Parent’s perceptions on important aspects of quality care by urbanicity

	Total Sample (n=295)	Rural (n=95)	Mid-sized (n=88)	Urban (n=112)
Providers are warm, caring, and provide empathy to students.	54.6%	56.80%	54.5%	52.7%
The program has trained educators who are qualified to work with the children.	33.6%	24.20%	38.6%	37.5%
Provides a stimulating environment for all children, including those with disabilities.	24.4%	26.30%	22.7%	24.1%
Teaching is interactive and engaging for the children.	46.8%	49.50%	50.0%	42.0%
Provides open and consistent communication to parents.	22.0%	25.30%	20.5%	20.5%
Uses a proven curriculum to maximize children’s learning and development.	20.7%	25.30%	22.7%	15.2%
Staffing patterns provide for adult supervision of children at all times.	11.2%	12.60%	9.1%	11.6%
Gathers information about each child to meet their individual needs.	16.6%	15.80%	14.8%	18.8%
Uses positive discipline and patiently guides the child’s behavior.	27.1%	29.50%	23.9%	27.7%
Provides careful supervision adjusted for different ages and abilities.	8.1%	11.60%	8.0%	5.4%
Provides a generous amount of supervised free playing time and social learning.	21.4%	16.80%	23.9%	23.2%
Education programs follow national health and safety best practices.	7.8%	4.20%	11.4%	8.0%
Offers small class sizes to maximize student to teacher ratios.	10.5%	8.5%	6.8%	15.2%

While there were few regional differences across the sample in terms of perceptions of quality care, slight differences emerged as a function of PTQ rating and auspice. For example, parents of children enrolled in LEA programs selected the option centered on warm and caring providers less frequently than parents of children enrolled in all other program types (while 62.1% of FCC parents selected this, only 39.3% of LEA parents did). The use of a proven curriculum was also selected as most important by 10.7% of LEA parents, compared to 33.3% of parents of children in other types of programs, while in total 20.7% of parents selected this option. Providing a

<sup>30</sup> 29 respondents selected zero options or more than 5, and their responses were filtered out.

generous amount of supervised free play time was also preferred differently by parents – while 21.4% of total parents selected this option, far more parents selected this if their child was enrolled in an FCC program (31.8%) compared to 16.7% of parents in other types of programs. Finally, small class sizes that maximize student-teacher ratios were also preferred differently - 10.5% of parents indicated this was important, while 25% of the LEA sample did.

Table 11. Parent’s perceptions on important aspects of quality care by auspice

	Total Sample (n=295)	FCC (n=66)	Center (n=85)	Ministry (n=104)	LEA (n=28)	Other (n=12)
Providers are warm, caring, and provide empathy to students.	<b>54.6%</b>	62.1%	60.0%	50.0%	39.3%	50.0%
The program has trained educators who are qualified to work with the children.	<b>33.6%</b>	22.7%	41.2%	33.7%	35.7%	33.3%
Provides a stimulating environment for all children, including those with disabilities.	<b>24.4%</b>	21.2%	31.8%	17.3%	32.1%	33.3%
Teaching is interactive and engaging for the children.	<b>46.8%</b>	48.5%	40.0%	51.0%	50.0%	41.7%
Provides open and consistent communication to parents.	<b>22.0%</b>	19.7%	18.8%	25.0%	21.4%	33.3%
Uses a proven curriculum to maximize children’s learning and development.	<b>20.7%</b>	12.1%	17.6%	29.8%	10.7%	33.3%
Staffing patterns provide for adult supervision of children at all times.	<b>11.2%</b>	15.2%	14.1%	7.7%	10.7%	0.0%
Gathers information about each child to meet their individual needs.	<b>16.6%</b>	9.1%	18.8%	15.4%	32.1%	16.7%
Uses positive discipline and patiently guides the child’s behavior.	<b>27.1%</b>	28.8%	27.1%	27.9%	21.4%	25.0%
Provides careful supervision adjusted for different ages and abilities.	<b>8.1%</b>	9.1%	7.1%	9.6%	3.6%	8.3%
Provides a generous amount of supervised free playing time and social learning.	<b>21.4%</b>	31.8%	18.8%	18.3%	17.9%	16.7%
Education programs follow national health and safety best practices.	<b>7.8%</b>	9.1%	2.4%	13.5%	3.6%	0.0%
Offers small class sizes to maximize student to teacher ratios.	<b>10.5%</b>	9.1%	8.2%	9.6%	25.0%	8.3%

There were also slight differences by PTQ rating. Although 20.7% of the sample selected using a proven curriculum as important, this was more important to parents of children in non-rated programs (32.5%), and less important to parents of children in 1- and 2-rated programs (11.9%). The focus on positive discipline to guide children’s behavior was also differently preferred: while 27.1% of the sample selected this, 38.1% of parents in 1- and 2-rated programs ranked this as important, compared to 20.0% of parents in non-rated and 0-rated programs. Small class sizes were ranked as important to 14.5% of parents in 4-rated programs, while none in 1- and 2-rated programs ranked this as important. Finally, a focus on health and safety behaviors was important to 7.8% of total parents, but 20% of parents in non-rated and 0-rated programs ranked this as important.

Table 12. Parent’s perceptions on important aspects of quality care by PTQ rating

	Total Sample (n=295)	Non-rated (n=40)	1 or 2 (n=42)	3 (n=130)	4 (n=83)
Providers are warm, caring, and provide empathy to students.	54.6%	50.0%	57.1%	53.8%	56.6%
The program has trained educators who are qualified to work with the children.	33.6%	35.0%	35.7%	29.2%	38.6%
Provides a stimulating environment for all children, including those with disabilities.	24.4%	25.0%	33.3%	24.6%	19.3%
Teaching is interactive and engaging for the children.	46.8%	50.0%	50.0%	47.7%	42.2%
Provides open and consistent communication to parents.	22.0%	25.0%	16.7%	23.1%	21.7%
Uses a proven curriculum to maximize children’s learning and development.	20.7%	32.5%	11.9%	18.5%	22.9%
Staffing patterns provide for adult supervision of children at all times.	11.2%	0.0%	7.1%	15.4%	12.0%
Gathers information about each child to meet their individual needs.	16.6%	15.0%	19.0%	14.6%	19.3%
Uses positive discipline and patiently guides the child’s behavior.	27.1%	20.0%	38.1%	28.5%	22.9%
Provides careful supervision adjusted for different ages and abilities.	8.1%	10.0%	4.8%	8.5%	8.4%
Provides a generous amount of supervised free playing time and social learning.	21.4%	20.0%	21.4%	22.3%	20.5%
Education programs follow national health and safety best practices.	7.8%	20.0%	4.8%	4.6%	8.4%
Offers small class sizes to maximize student to teacher ratios.	10.5%	5.0%	0.0 %	13.1%	14.5%

In addition, we asked parents how confident they feel that they fully understand what constitutes a high-quality environment, rating their choice from 1 (not confident) to 7 (very confident). Parents feel very confident that they understand what constitutes a high-quality environment - 49.2% of parents feel very confident, and just over three-fourths of the sample chose 6 or 7 as their response to this question. These levels of confidence varied somewhat based on the programs parents chose to enroll their children in. For example, 34.9% of parents in non-rated programs selected “very confident” as their response, compared to 40% in 1- and 2-rated programs, 52.9% in 3-rated programs, and 55.6% in 4-rated programs. Parents of pre-K-aged children were also more likely to rate themselves “very confident” (51.6%) compared to parents of infants/toddlers (43.8%). Finally, parents of children in center-licensed programs rated themselves very confident more often (54.2%) than FCCs (47.1%), LEAs (50%), ministry programs (47.3%), and other programs (38.5%).

Lastly, we asked parents to respond to the question: “In your own words: What is a high-quality early care and education environment?” A total of 287 parents provided a description of quality in ECE. An analysis of the responses showed a few key themes. Almost 70% of respondents mentioned something about academics, learning, school readiness, or preparation for

kindergarten in their write-up. Some of these responses focused solely on academics (e.g., *“Being able to start kindergarten knowing they are ready because of the skills they have learn in daycare,”* and *“Places emphasis on learning.”*). Others highlighted academic readiness but also wanted their child’s ECE program to focus on their non-academic skills as well (e.g., *“One that supports the child’s learning and development both social and academic, but also allows them to feel safe and cared for during the day,”* and *“Somewhere where the parent and teacher work together to be on the same page and teach social and emotional learning along with fundamental basics.”*). Much like in the quantitative responses, a majority of parents/caregivers also mentioned something about their child’s teacher being warm, nurturing, caring, positive, or that their child would feel comfortable at their ECE program. For instance, parents/caregivers mentioned *“An environment in which the teacher is actively engaging with my child to advance her knowledge and skills that are age appropriate while maintaining a safe/nurturing atmosphere”* and *“A facility with educated + compassionate staff that has the ability to teach + expose a child to developmentally appropriate skills AND meets our exceed NAEYC expectations.”*

In 84 of the responses (29% of parents), a focus on safety was highlighted as important (e.g., *“Something that is healthy and safe for my child,”* and *“Safe, disciplined, active in on going learning.”*). Another theme that came up frequently in the responses (27% of respondents/78 caregivers) was a focus on fun: The respondents mentioned words like *“play,” “fun,” “play-based learning,”* or described the types of toys children would have access to. Samples of responses that matched this theme include: *“When they put learning and fun for your children together. When they bring progress home,”* and *“Environment with stimulating activities leading to educational growth while also giving children freedom to play.”* Finally, parents/caregivers were also interested in children having the opportunity to gain social skills or socialize with similar-age or different-age peers. In total, 81 respondents (28%) mentioned these types of opportunities, with responses ranging from *“Teaching kids to be well rounded; educationally, socially, and behaviorally,”* to *“A center that provides academic learning while also focusing on the social skills needed at this young age.”* Other themes that came up less frequently included a focus on socio-emotional skills (16%), a focus on children growing/maturing in general (10%), having skilled or highly qualified/educated teachers (9%), meeting the needs of the caregiver’s individual child/being able to be inclusive and meet the needs of all children (9%), communicating well and/or meeting the needs of families (8%), providing an environment that is clean (7%), providing developmentally appropriate or age appropriate teaching and caregiving (6%), providing an environment that encourages children to be curious and explore (5%), and one that has a high Paths to Quality rating or follows the guidelines of Paths to Quality (2%).



Figure 29. Parent’s definitions of a high-quality early care and education environment



## Discussion of Findings

This second report summarizes findings for the 2021 and 2022 school years for a landscape of ECE programs in the state of Indiana. The first report focused on information on program quality and program characteristics. This second report looks at another year of data on quality, and incorporates the findings on teacher self-reported quality, child growth across programs, and parental resources and engagement, as well parental perceptions of various aspects of programs and quality.

The study looked at a range of program types for infants, toddlers, and pre-K classrooms, and their observed quality levels vary as measured by the CLASS tool, as well as by teacher self-reported tools. Across all ages, programs were found to provide nurturing and safe environments for children. However, findings show that scores were lower for aspects related to language development and facilitated instruction. The results of the study suggest that teachers who work with infants, toddlers, and preschool-age children could benefit significantly from targeted support in several key areas. Specifically, teachers need assistance in promoting language use, scaffolding learning, linking concepts across activities, and encouraging metacognition in their students.

Effective language development is critical to young children's cognitive and social-emotional growth, and children benefit from ample opportunities for children to engage in language-rich environments. Scaffolding learning involves providing appropriate support and guidance to help children build on their existing knowledge and skills, and teachers can benefit from training in how to scaffold learning experiences intentionally and effectively. Linking concepts across activities is another critical area where teachers could benefit from targeted support. Children's

learning is most effective when they can connect what they learn in one activity to other experiences in their daily lives. Cross-content activities that allow children to apply their learning in diverse contexts and areas are examples of such approaches.

Finally, promoting metacognition involves helping children understand how they learn. Teachers should encourage children to reflect on their learning experiences and to monitor their progress, which can enhance their confidence and independence. Therefore, providing intentional support in these key areas can help teachers provide the best learning experiences for their students.

Based on program region, auspice, and PTQ rating, some differences were observed. In particular, 3- and 4-rated programs on average tended to score higher on classroom quality. An evaluation of the PTQ system in toddler and preschool classrooms conducted by Purdue University documented similar results. CLASS scores were higher for 3- and 4-rated programs, and lower scores in the language and instructional support-related domains were observed. (Elicker et al., 2018) These findings suggest that while the PTQ rating system captures meaningful differences in program quality, many programs struggle to provide high-level instructional experiences in the classroom. Given the timing of this study in relation to the COVID-19 pandemic, and the consistency of the findings in quality across two years, the issues on quality in these areas seem to transcend the impact of the pandemic on the early childhood sector.

Infants and toddlers in the sample showed lower developmental levels than typically observed for children in this age group (with the measures utilized in this study). However, babies and toddlers appear to be gaining skills relative to the norm, with positive gains in standardized scores. Some differences emerged by program type, with children in programs showing lower baseline scores in language and cognitions seemingly evidencing higher growth.

When the study commenced, the children of preschool age in the sample had lower literacy and math skills compared to typical peers (children with which the instruments were normed), but their language and executive function skills were similar. Over the course of the study, children made on average significant gains in receptive vocabulary, exceeding the typical growth for children their age. In addition, they exhibited mostly expected growth in math and executive functions. However, their literacy skills showed slower growth than typical, and in fact, a decline was observed relative to their peers in standard scores. These patterns were consistent across different ratings, urban areas, and program types. The slower growth in literacy development is particularly significant since children were lagging to start with, and these trends mean a slower literacy growth is compounding over time. The math trends are not as stark, but there is some indication of a similar (albeit smaller) issue for some children. These trends seem to align with other research on the COVID-19 pandemic's impact on children's learning; for instance, researchers from NWEA found that second through fifth graders showed similar growth trends on MAP math and reading tests in 2022 as compared to pre-pandemic years; however, first graders (who had only experienced pandemic schooling) showed 6-7% lower growth than would be expected in a typical year. (Kuhfeld & Lewis, 2022) It could be that children in the critical early years evidence higher vulnerability to the impacts of the pandemic (as suggested by this research and other studies measuring child growth during the pandemic); further research could



help clarify these and other similar findings. Feasible, an emphasis in vocabulary over literacy and math across all types of programs could be driving the observed trends.

Finally, parents felt fairly confident that they understand what constitutes a high-quality early learning environment. In both qualitative and quantitative feedback, they focused on warm and nurturing providers as critical components of a high-quality early learning setting; they focused also on interactive and engaging teaching, and that academic skills are focused on. In addition, parents reported frequently engaging in the types of activities that support children's early learning, such as reading books with them and teaching them about the world around them.

## Acknowledgments

We are thankful to Early Learning Indiana, with generous support from Lilly Endowment Inc., for engaging with NIEER for this work. We are also thankful to the Transform Consulting Group for their partnership. Finally, we are thankful to all the program directors and teachers that opened their classrooms to the research group. This includes the partnership with Building Blocks of Indiana (formerly Community Coordinated Childcare Inc (4Cs) of Indiana) for access to their data.

## References

- Bandel, E., Aikens, N., Vogel, C. A., Boller, K., & Murphy, L. (2014). Observed quality and psychometric properties of the CLASS-T in the early head start family and child experiences survey. Mathematica Policy Research.
- Barnett, W. S., & Friedman-Krauss, A. H. (2016). *State(s) of Head Start*. New Brunswick, NJ: National Institute for Early Education Research.
- Bassok, D., Magouirk, P., & Markowitz, A. J. (2021). Systemwide quality improvement in early childhood education: Evidence from Louisiana. *AERA Open*, 7, 23328584211011610.
- Bichay-Awadalla, K. & Bulotsky-Shearer, R.J. (2022) Examining the Factor Structure of the Classroom Assessment Scoring System Toddler (CLASS-T) in Early Head Start and Subsidized Child Care Classrooms, *Early Education and Development*, 33:2, 309-325, DOI: 10.1080/10409289.2021.1887700
- Decker-Woodrow, L., Hartman, N., Diaz, E., Adachi, E., Lamey, G. & Barfield, D. (2019). Pre-K 4 SA Evaluation Report. Year 6. San Antonio, TX: Westat.
- Elicker, J., Lane, S., Gold, Z. S., Mishra, A., & Christ, S. (2018). Final report: Paths to QUALITY evaluation. Center for Families Publications.  
<https://docs.lib.purdue.edu/cffpub>
- Finders, J. K., Budrevich, A., Duncan, R. J., Purpura, D. J., Elicker, J., & Schmitt, S. A. (2021). Variability in Preschool CLASS Scores and Children’s School Readiness. *AERA Open*, 7, 23328584211038938.
- Hamre, B. K., Karen, M., Paro, L., Pianta, R. C., & LoCasale-Crouch, J. (2014). *Classroom assessment scoring system (CLASS) manual, infant*. Paul H. Brookes Publishing.
- Head Start, ECLKC (2020) A National Overview of Grantee CLASS® Scores in 2020. Washington, DC: Author. <https://eclkc.ohs.acf.hhs.gov/data-ongoing-monitoring/article/national-overview-grantee-class-scores-2020>
- Kuhfeld, M., & Lewis, K. (2022). *COVID-19 in the early elementary years: A comparison of achievement in spring 2019 and spring 2022*. Collaborative for Student Growth. NWEA Research.
- La Paro, K. M., Hamre, B. K., & Pianta, R. C. (2012). *Classroom assessment scoring system (CLASS) manual, toddler*. Paul H. Brookes Publishing Company.
- LA Department of Education (undated). Infant CLASS Lessons from the Pilot Year. LA: Author. [https://louisianabelieves.com/docs/default-source/teacher-leader-summit/2019-teacher-leader-summit/e054---infant-class-slides.pdf?sfvrsn=1d0b9c1f\\_7](https://louisianabelieves.com/docs/default-source/teacher-leader-summit/2019-teacher-leader-summit/e054---infant-class-slides.pdf?sfvrsn=1d0b9c1f_7)
- NIEER (2014). The Evaluation of Quality of New Jersey’s Preschool Classrooms 2014. Retrieved from <https://nieer.org/wp-content/uploads/2021/01/The-Evaluation-of-Quality-of-NJs-Preschool-Classrooms-2014.pdf>
- Nores, M., Barnett, W.S., Jung, K., Joseph, G. & Bachman, L. (2019). Year 4 report: Seattle Preschool Program evaluation. New Brunswick, NJ: National Institute for Early Education Research & Seattle, WA: Cultivate Learning, 66 pp. <http://nieer.org/research-report/seattle-pre-k-program-evaluation>

- Nores, M., Valle, E., Contreras, C. & M. Allenger (2020). Evaluation of West Virginia Universal Pre-K. Classroom observation findings. New Brunswick, NJ: National Institute for Early Education Research.
- Nores, M., Harmeyer, E., Connors-Tadros, L., Li, Z. & Contreras, C. (2022). Evaluation of the Early Childhood Programs in Indiana. Report 1. New Brunswick, NJ: National Institute for Early Education Research.
- NYC Department of Education (2021) Early Childhood Program Assessments Classroom Assessment Scoring System (CLASS) and Early Childhood Environmental Rating Scale Revised (ECERS-R) Release. New York: Author.
- Pianta, R. C., La Paro, K. M., & Hamre, B. K. (2012). *Classroom assessment scoring system (CLASS). Manual*. pre-K. Paul H. Brookes Publishing Company.
- Weiland, C., Greenberg, E., Bassok, D., Markowitz, A., Guerrero Rosada, P., Luetmer, G., Abenavoli, R., Gomez, C., Johnson, A., Jones-Harden, B., Maier, M., McCormick, M., Morris, P., Nores, M., Phillips, D., & Snow, C. (2021). Historic crisis, historic opportunity: Using evidence to mitigate the effects of the COVID-19 crisis on young children and early care and education programs. *University of Michigan Education Policy Initiative and Urban Institute Policy Brief*, <https://edpolicy.umich.edu/files/EPI-UI-Covid%20Synthesis%20Brief%20June%202021.pdf>
- 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.
- Xue, Y., Baxter, C., Jones, C., Shah, H., Caronongan, P., Aikens, N., & Atkins-Burnett, S. (2021). Early Head Start Programs, Staff, and Infants/Toddlers and Families Served: Baby FACES 2018 Data Tables. OPRE Report 2021-92. Administration for Children & Families. <https://files.eric.ed.gov/fulltext/ED613544.pdf>
- Xue, Y., Del Grosso, P., & Carlson, B. A Snapshot of Quality in Child Care Centers That Partner with Early Head Start Programs: Insights from Baby FACES 2018 (No. 712ce2db10a64f818994c28bd7e0d5d2). Mathematica Policy Research. [https://www.acf.hhs.gov/sites/default/files/documents/opre/ehs\\_ccp\\_quality\\_brief\\_june2022.pdf](https://www.acf.hhs.gov/sites/default/files/documents/opre/ehs_ccp_quality_brief_june2022.pdf)

## Appendix A. Measures

### Classroom Observation Measures

#### *Classroom Assessment Scoring System (CLASS Infant; Hamre et al., 2014)*

The Classroom Assessment Scoring System (CLASS) Infant is an observational tool used to measure the quality of interactions between teachers and children ages 0 through 18 months. Teachers are assessed on their interactions in one domain - Responsive Caregiving (RC), which consists of four dimensions: Relational Climate, Teacher Sensitivity, Facilitated Exploration, and Early Language Support. Observers complete five 15-minute cycles, and code their observations for 10 minutes each between cycles. Observers assign scores on a seven-point Likert-type scale. Scores of 1-2 indicate low quality, scores of 3-5 are in the moderate range, and scores of 6-7 indicate high quality.

#### *Classroom Assessment Scoring System (CLASS Toddler; La Paro et al., 2012)*

The Classroom Assessment Scoring System (CLASS) Toddler is an observational tool used to measure the quality of interactions between teachers and children ages 15 months through 36 months. CLASS Toddler is divided into two broad domains that cover eight dimensions. The Emotional and Behavioral Support (EBS) domain measures the social and emotional supports provided by teachers and how teachers manage children's time, behavior, and attention in the classroom. The dimensions included are: Positive Climate, Negative Climate, Teacher Sensitivity, Regard for Child Perspectives, and Behavior Guidance. The Engaged Support for Learning (ESL) domain measures how well teachers promote cognitive and language development. The dimensions covered include Facilitation of Learning and Development, Quality of Feedback, and Language Modeling.

#### *Classroom Assessment Scoring System (CLASS Pre-K; Pianta, La Paro, & Hamre, 2008; Pianta & Hamre, 2009; Hamre et al., 2014)*

The Classroom Assessment Scoring System (CLASS) is an observational system that assesses classroom practices by measuring the interactions between students and teachers. CLASS measures interactions along ten distinct dimensions, which are grouped into three overarching domains. The Emotional Support (ES) domain is measured by four dimensions: Positive Climate, Negative Climate, Teacher Sensitivity, and Regard for Student Perspectives. The Classroom Organization (CO) domain is measured by three dimensions: Productivity, Behavior Management, and Instructional Learning Formats. The Instructional Support (IS) domain is measured by three dimensions: Concept Development, Quality of Feedback, and Language Modeling. Observations consist of five 20-minute cycles, with 10-minute coding periods between each cycle. Scores (codes) are assigned during various classroom activities and then averaged across all cycles for overall scores in three domains. Each dimension is scored on a seven-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.

Table A.1. CLASS Infant Domains and Dimension Descriptions.

Domain	Dimension	Description
<b>Responsive Caregiving</b>	Relational Climate	Assesses the emotional connection, respect, and enjoyment demonstrated between teachers and how infants respond to these connections.
	Teacher Sensitivity	Focuses on teachers' awareness of and responsivity to children's cues, including verbal and nonverbal, and ability to provide comfort in a timely manner.
	Facilitated Exploration	Captures the degree to which the teachers' interactions with children during routine care and playtime support their development and engagement.
	Early Language Support	Measures the extent to which teachers facilitate and encourage children in language use, including language-stimulation and language-facilitation techniques.

Table A.2. CLASS Toddler Domains and Dimension Descriptions.

Domain	Dimension	Description
<b>Emotional and Behavioral Support</b>	Positive Climate	Measures the warmth of relationships between teachers and children, and the ways teachers demonstrate respect for children.
	Negative Climate	Assesses the level of expressed negativity such as anger, hostility, or aggression exhibited by teachers and/or students in the classroom.
	Teacher Sensitivity	Captures teachers' awareness of and responsivity to children, including their body language and behavior.
	Regard for Student Perspectives	Measures how well teachers emphasize children's motivations, points of view, and interests, and how much they maximize children's independence.
	Behavior Guidance	Captures how effectively teachers monitor, prevent, and redirect behavior.
<b>Engaged Support for Learning</b>	Facilitation of Learning and Development	Emphasizes how teachers facilitate instruction in a way that allows children to take an active role in their learning, and how well they connect children's lived experiences with classroom content.
	Quality of Feedback	Focuses on how well teachers extend students' learning through providing specific feedback or scaffolding and how well they encourage and affirm children.
	Language Modeling	Measures the extent to which teachers respond to and extend upon children's use of language.

Table A.3. CLASS Pre-K Domains and Dimension Descriptions.

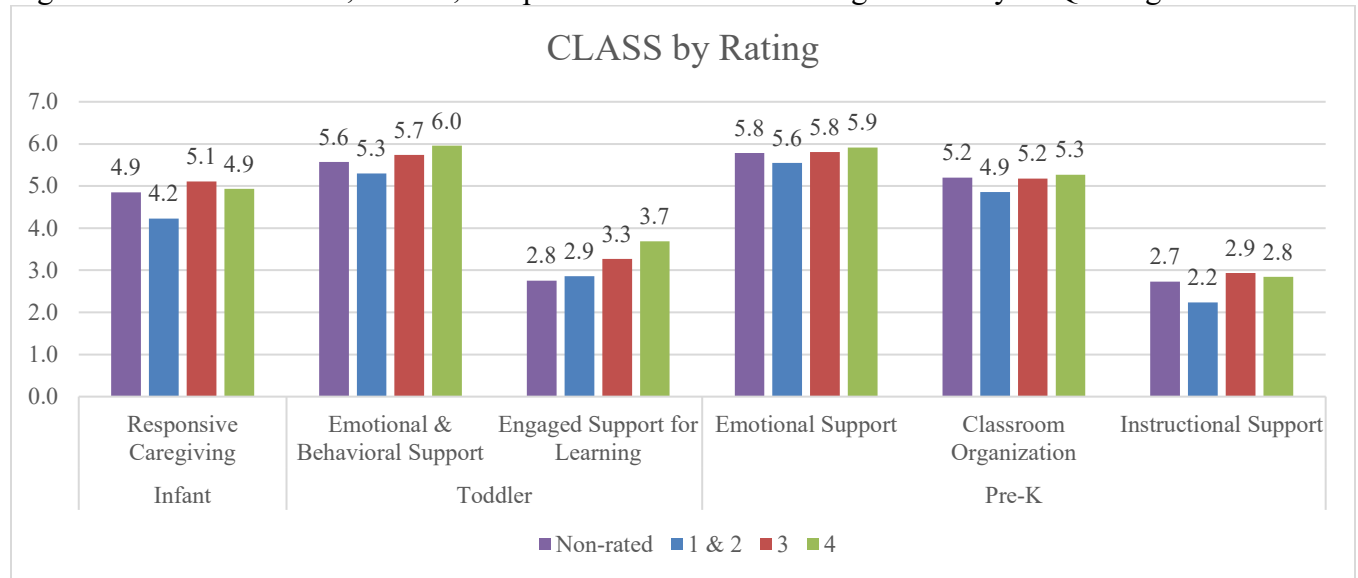
Domain	Dimension	Description
<b>Emotional Support</b>	Positive Climate	Reflects the emotional connection between teachers and children and among children, and the warmth, respect, and enjoyment communicated by verbal and nonverbal interactions.
	Negative Climate	Reflects the overall level of expressed 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.

## Evaluation Report: ECE Programs in Indiana

<b>Classroom Organization</b>	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.
<b>Instructional Support</b>	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.

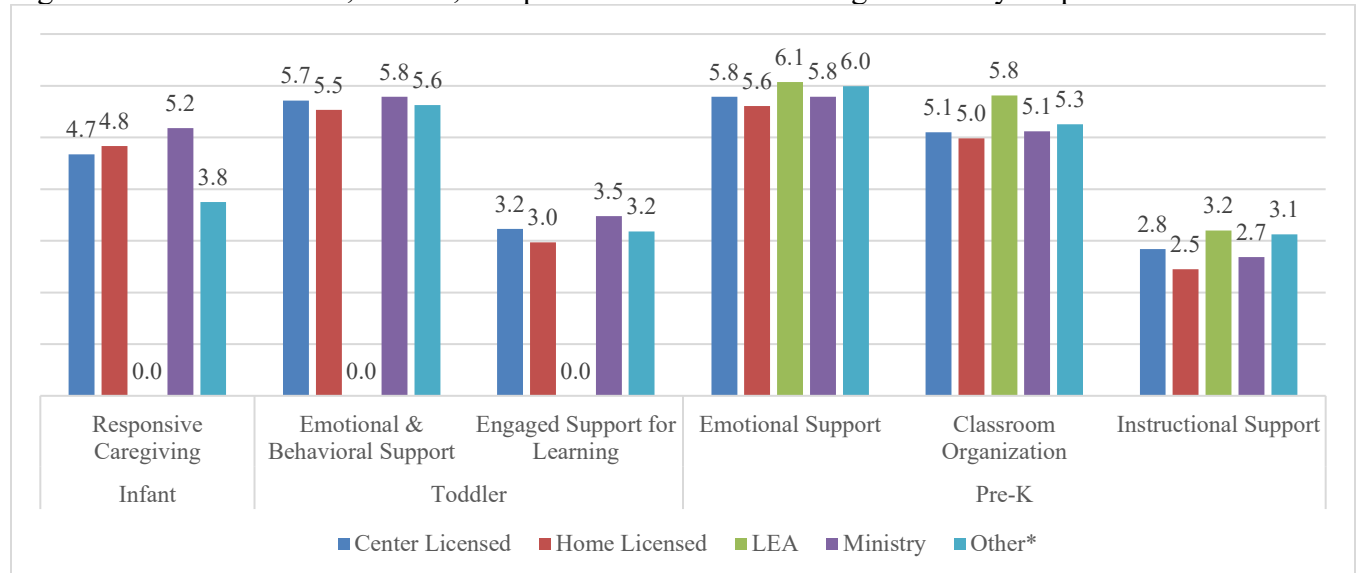
## Appendix B. Findings on Classroom Observations.

Figure B.1. CLASS infant, toddler, and pre-K 2021 & 2022 average scores by PTQ ratings



Note: No statistical difference is evidenced on any of the CLASS scales.

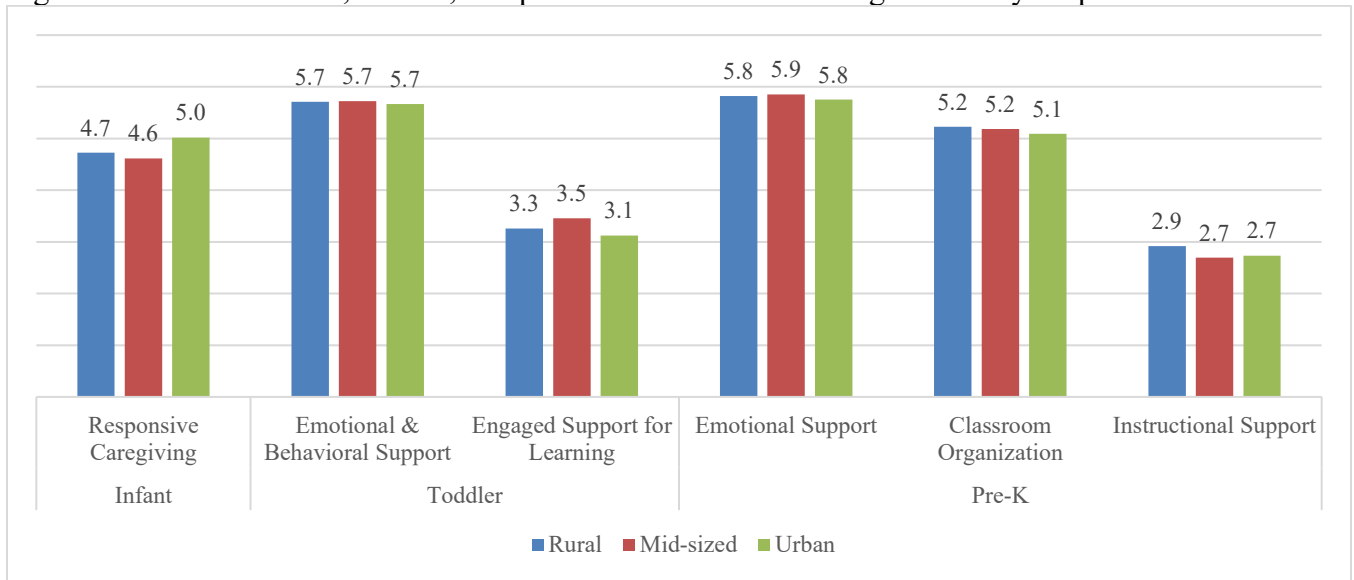
Figure B.2. CLASS infant, toddler, and pre-K 2021 & 2022 average scores by auspice



Note: Differences are statistically significant for: LEA facilities scored higher on Classroom Organization and Instruction Support scales than Center Licensed, Home Licensed, and Ministry facilities.



Figure B.3. CLASS infant, toddler, and pre-K 2021 and 2022 average scores by auspice



Note: No statistical difference is evidenced on any of the CLASS scales.

## Appendix C. Findings on Baseline Child Assessments

Table C.1. Bayley cognitive gains by child and center characteristics

		COG Raw Score Gains			COG Standard Score Gains		
		Valid N	Mean	St. Dev.	Valid N	Mean	St. Dev.
Total		178	16.37	18.90	178	2.50	17.04
Gender	Male	89	15.67	22.26	89	1.35	18.57
	Female	89	17.07	14.89	89	3.65	15.39
Age	1	34	19.09	15.42	34	1.03	22.35
	2	72	16.28	19.24	72	2.08	16.84
	3	72	15.18	20.11	72	3.61	14.35
Race	White	98	18.96	21.25	97	4.79	17.15
	Non-white	23	10.22	11.04	23	-2.17	16.36
	Missing	57	14.40	16.34	58	0.52	16.77
Facility Type	Center Licensed	68	15.04	16.53	69	2.39	16.90
	Home Licensed	39	20.82	22.90	38	7.76	17.11
	Ministry	63	15.05	18.86	63	-0.71	16.65
	Other	8	16.38	16.50	8	3.75	17.88
Facility Quality	1 & 2	45	19.20	17.55	45	6.22	15.71
	3	52	21.02	24.49	52	5.77	17.05
	4	57	12.91	14.94	57	-0.61	18.10
Region	Non-rated	24	9.21	11.69	24	-4.17	14.12
	Mid-sized	47	16.68	21.95	47	2.34	17.69
	Rural	55	21.71	19.51	55	7.82	16.15
	Urban	76	12.32	15.35	76	-1.25	16.45

Note: Group differences are statistically significant for raw and standard gains for the following: children from Non-rated and Level 4 facilities gained less than that from Level 1 & 2 and Level 3 facilities on raw and standard scores for Bayley Cognitive scale; children from rural region gained more than from urban region on raw and standard scores for Bayley Cognitive scale.

Table C.2. Bayley language gains by child and center characteristics

		LANG Raw Score Gains			LANG Standard Score Gains		
		Valid N	Mean	St. Dev.	Valid N	Mean	St. Dev.
Total		178	8.86	11.83	176	2.98	18.63
Gender	Male	89	8.15	12.18	88	0.97	17.92
	Female	89	9.57	11.49	88	5.00	19.20
Age	1	34	5.21	6.10	34	-2.41	15.09
	2	72	8.58	12.33	71	2.97	20.92
	3	72	10.86	13.01	71	5.58	17.37
Race	White	98	10.40	13.70	97	4.92	20.97
	Non-white	23	10.00	9.53	22	3.00	16.64
	Missing	57	5.75	8.18	57	-0.32	14.47
Facility Type	Center Licensed	68	8.91	11.78	68	4.90	17.95
	Home Licensed	39	11.08	10.62	37	8.76	18.10
	Ministry	63	8.16	12.48	63	-2.14	19.09
	Other	8	3.12	12.18	8	0.38	14.90
Facility Quality	1 & 2	45	8.00	11.33	44	3.27	17.02
	3	52	11.9	12.16	52	7.35	19.63
	4	57	8.79	11.38	56	2.45	17.07
	Non-rated	24	4.04	11.90	24	-5.75	20.58
Region	Mid-sized	47	9.70	12.38	46	6.59	18.96
	Rural	55	8.00	11.94	55	1.38	18.31
	Urban	76	8.96	11.51	75	1.95	18.61

Note: Group differences are statistically significant for raw and standard gains for the following: children from Ministry facilities gained less than those from Center Licensed and Home Licensed facilities on both raw and standard scores for Bayley Language; children from Non-rated facilities gained more less those from Level 3 facilities for both raw and standard scores.

Table C.3. Bayley social-emotional gains by child and center characteristics

		SOEM Raw Score Gains			SOEM Standard Score Gains		
		Valid N	Mean	St. Dev.	Valid N	Mean	St. Dev.
Total		101	0.67	5.38	101	3.37	26.91
Gender	Male	49	-0.76	5.48	49	-3.78	27.42
	Female	52	2.02	4.97	52	10.1	24.84
Age	1	21	-1.19	7.29	21	-5.95	36.46
	2	42	1.10	4.88	42	5.48	24.39
	3	38	1.24	4.55	38	6.18	22.73
Race	White	75	0.12	5.65	75	0.60	28.27
	Non-white	9	1.67	3.32	9	8.33	16.58
	Missing	17	2.59	4.66	17	12.94	23.32
Facility Type	Center Licensed	23	-2.65	6.05	23	-13.26	30.25
	Home Licensed	23	1.35	3.55	23	6.74	17.75
	Ministry	45	1.16	4.77	45	5.78	23.86
	Other	10	4.60	6.52	10	23.00	32.59
Facility Quality	1 & 2	29	0.14	7.20	29	0.69	36.02
	3	27	-0.11	4.88	27	-0.56	24.39
	4	26	1.19	3.92	26	5.96	19.60
Region	Non-rated	19	1.89	4.59	19	9.47	22.97
	Mid-sized	18	-0.39	3.63	18	-1.94	18.16
	Rural	44	0.00	6.35	44	0.00	31.77
	Urban	39	1.92	4.69	39	9.62	23.43

Note: Group differences are statistically significant for raw and standard gains between boys and girls, between children from Center-licensed facilities and those from all other types of facilities.

Table C.4. Bayley socialization gains by child and center characteristics

		SOC Raw Score Gains			SOC Standard Score Gains		
		Valid N	Mean	St. Dev.	Valid N	Mean	St. Dev.
Total		88	0.36	6.19	88	0.89	16.77
Gender	Male	43	-0.21	6.49	43	-0.81	17.55
	Female	45	0.91	5.90	45	2.51	16.02
Age	1	21	-1.10	6.85	21	-3.90	18.50
	2	33	1.09	5.90	33	3.18	16.19
	3	34	0.56	6.06	34	1.62	16.11
Race	White	67	0.73	6.15	67	1.78	16.70
	Non-white	8	-1.75	8.03	8	-4.25	21.37
	Missing	13	-0.23	5.25	13	-0.54	14.64
Facility Type	Center Licensed	22	-0.86	7.54	22	-2.64	20.65
	Home Licensed	20	-0.35	6.88	20	-0.90	18.03
	Ministry	40	0.80	4.93	40	2.08	13.57
	Other	6	4.33	5.35	6	11.83	14.40
Facility Quality	1 & 2	28	0.93	8.14	28	2.43	22.00
	3	23	-1.22	5.55	23	-3.52	15.05
	4	25	1.60	4.57	25	4.20	12.42
Region	Non-rated	12	-0.50	4.72	12	-1.17	13.03
	Mid-sized	15	0.27	6.57	15	1.07	18.36
	Rural	43	0.93	5.96	43	2.07	16.12
	Urban	30	-0.40	6.43	30	-0.90	17.30

Note: No group differences are statistically significant for raw and standard gains.

Table C.5. PPVT IV score gains by child and center characteristics

		PPVT Raw Score Gains			PPVT Standard Score Gains		
		Valid N	Mean	St. Dev.	Valid N	Mean	St. Dev.
Total		329	7.42	12.50	506	17.04	15.51
Gender	Female	173	6.17	12.66	266	16.17	15.94
	Male	156	8.81	12.22	240	18.01	19.78
Age	3 and younger	90	8.92	11.26	136	11.79	16.65
	4	187	6.70	12.89	286	17.98	16.93
	5 and older	52	7.42	13.14	84	22.35	20.27
Ethnicity	African American	54	7.24	14.38	56	22.61	19.38
	Hispanic/Latino	23	9.70	9.92	23	28.57	22.33
	Missing	18	10.50	10.82	181	6.07	12.34
	Other	15	6.20	9.99	15	28.00	12.69
	White/Non-Hispanic	219	7.06	12.57	231	22.43	17.63
Language	Missing	3	6.33	5.69	165	4.19	11.93
	English	314	7.33	12.63	329	23.15	16.60
Dev. Concern	Non-English	12	10.17	10.29	12	26.33	31.11
	Missing	1	17.00	.	163	4.07	10.35
Facility Type	No	315	7.52	12.26	329	23.24	16.51
	Yes	13	4.31	17.97	14	22.36	35.38
	Center Licensed	108	7.19	12.91	192	16.59	18.76
Facility Quality	Home Licensed	63	6.24	12.75	67	18.69	17.71
	LEA	39	6.64	12.34	52	20.81	18.79
	Ministry	96	8.42	12.33	153	16.04	16.59
	Other	23	8.96	11.44	42	15.48	17.66
Region	1 & 2	63	7.44	13.41	69	22.72	15.48
	3	137	8.14	12.60	232	16.05	17.50
	4	84	6.23	11.88	132	16.76	19.69
	Non-rated	45	7.44	12.27	73	15.34	16.98
Community Poverty	Mid-sized	84	9.42	12.75	145	16.89	16.66
	Rural	81	6.48	12.34	129	15.12	17.15
Community Poverty	Urban	164	6.87	12.42	232	18.20	18.98
	Low	248	7.13	12.53	390	16.66	17.74
Community Poverty	High	81	8.31	12.46	116	18.31	18.36

Note: Group differences are statistically significant for standard gains between all age cohorts (age 3 gained less than age 4 and 5), ethnicity groups (Missing groups gained lower than all other groups), home language groups (missing group gained less than other two groups) developmental concerns categories (missing group gained less than other two groups), facility quality levels (Level 1 & 2 gained more than other level categories).

Figure C.1. PPVT IV standard score gains by center characteristics

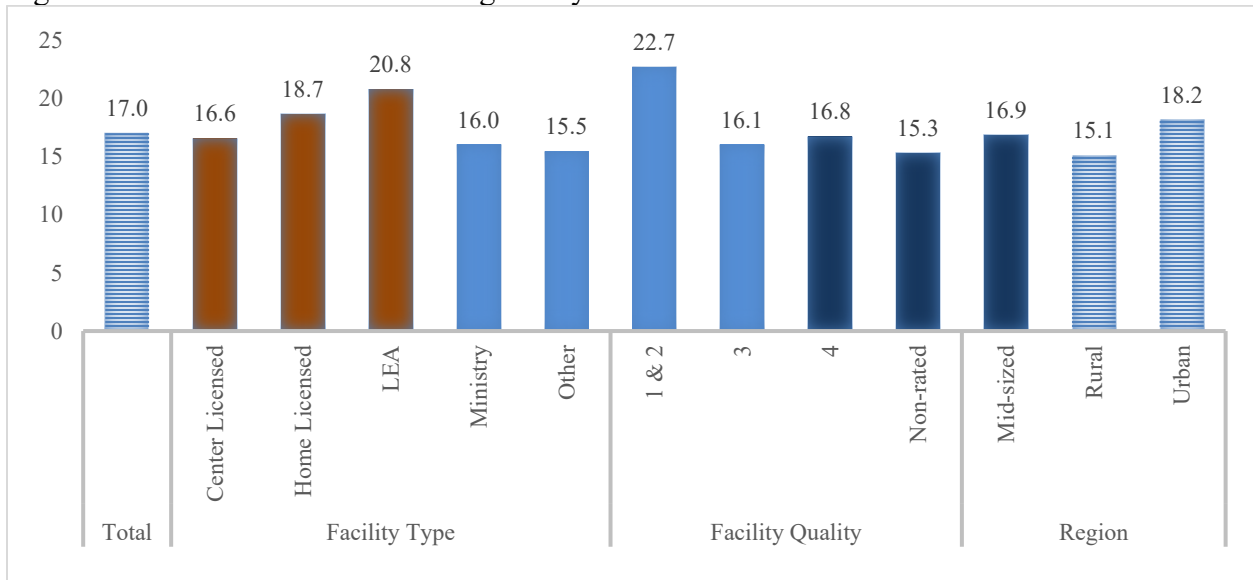


Table C.6. WJ IV-LW score gains by child and center characteristics

		LW Raw Score Gains			LW Standard Score Gains		
		Valid N	Mean	St. Dev.	Valid N	Mean	St. Dev.
Total		487	1.62	2.11	482	-1.79	6.46
Gender	Female	254	1.82	2.23	249	-1.24	6.46
	Male	233	1.39	1.96	233	-2.39	6.43
Age	3 and younger	132	1.08	1.97	125	-2.66	6.44
	4	277	1.81	2.10	277	-1.47	6.47
	5 and older	78	1.85	2.25	80	-1.56	6.40
Ethnicity	African American	55	1.16	1.78	56	-3.21	6.26
	Hispanic/Latino	23	2.09	2.02	22	-1.86	5.49
	Missing	170	1.79	1.99	170	-0.60	6.48
	Other	14	0.86	2.03	14	-4.64	4.70
	White/Non-Hispanic	225	1.60	2.27	220	-2.16	6.55
Language	Missing	154	1.75	1.94	156	-0.44	6.61
	English	322	1.55	2.20	315	-2.41	6.33
Dev. Concern	Non-English	11	1.64	1.96	11	-3.18	5.40
	Missing	152	1.77	1.94	154	-0.36	6.60
	No	321	1.56	2.19	314	-2.44	6.31
Facility Type	Yes	14	1.29	2.02	14	-3.00	6.14
	Center Licensed	186	1.33	1.96	184	-2.03	6.52
	Home Licensed	66	1.11	2.08	63	-4.22	6.62
	LEA	48	2.31	2.06	50	-0.20	6.18
	Ministry	145	1.94	2.27	143	-1.17	6.28
Facility Quality	Other	42	1.79	1.98	42	-1.14	6.12
	1 & 2	64	1.27	1.96	65	-3.09	6.02
	3	219	1.51	2.15	216	-2.13	6.42
	4	132	1.89	2.07	131	-1.25	6.66
Region	Non-rated	72	1.75	2.18	70	-0.56	6.41
	Mid-sized	141	1.38	2.18	139	-2.06	6.94
	Rural	123	1.67	2.10	122	-1.07	6.27
Community Poverty	Urban	223	1.74	2.07	221	-2.02	6.24
	Low	378	1.56	2.12	375	-1.95	6.68
	High	109	1.80	2.09	107	-1.22	5.62

Note: Group differences are statistically significant for raw score gains between boys and girls, age groups (age 3 gained less than age 4 and 5); facility types (Center-licensed and Home-licensed facilities gained less LEA and Ministry facilities). Group differences are statistically significant for standard score gains between ethnicity groups (missing group gained greater than African American, Other, and White), home language groups (missing group gained greater than English group), developmental concern groups (missing group gained greater than other two groups), facility types (Home-licensed center gained less than all other types of facilities).



Figure C.2. WJ IV-LW standard score gains by center characteristics

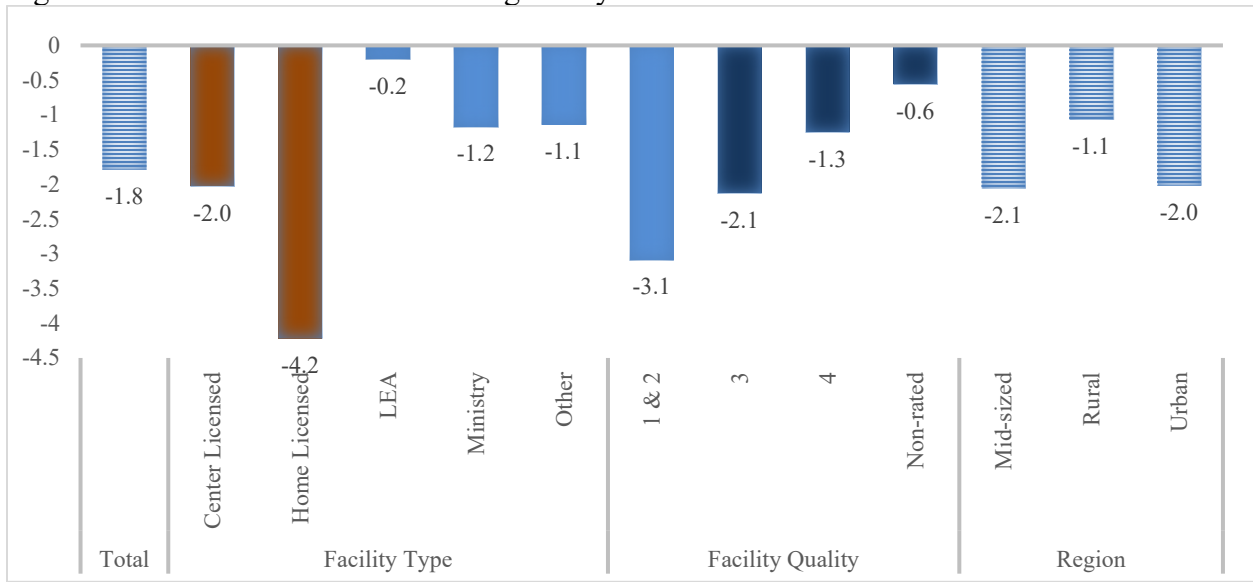


Table C.7. WJ IV-AP score gains by child and center characteristics

		AP Raw Score Gains			AP Standard Score Gains		
		Valid N	Mean	St. Dev.	Valid N	Mean	St. Dev.
Total		504	1.68	2.70	499	-0.36	10.79
Gender	Female	264	1.70	2.66	264	-0.21	10.90
	Male	240	1.66	2.75	235	-0.54	10.69
Age	3 and younger	137	2.23	2.60	132	1.58	10.48
	4	287	1.55	2.78	286	-0.55	10.98
	5 and older	80	1.19	2.47	81	-2.86	10.19
	Missing	57	1.30	2.78	54	-2.15	10.60
Ethnicity	African American	23	1.61	2.39	22	-1.45	11.00
	Hispanic/Latino	182	2.19	2.61	179	1.91	10.18
	Other	15	1.47	2.56	15	-1.33	9.29
	White/Non-Hispanic	227	1.38	2.75	229	-1.55	11.16
Language	Missing	167	2.14	2.62	164	1.71	10.29
	English	325	1.42	2.74	324	-1.44	10.90
	Non-English	12	2.33	1.78	11	0.55	11.39
Developmental Concern	Missing	165	2.17	2.65	162	1.85	10.37
	No	325	1.45	2.72	323	-1.46	10.96
Facility Type	Yes	14	1.14	2.25	14	-0.64	8.05
	Center Licensed	190	1.64	2.55	191	0.00	10.52
	Home Licensed	66	1.58	2.86	67	-2.22	11.97
Facility Quality	LEA	52	1.56	2.79	51	-0.78	10.53
	Ministry	155	1.95	2.88	149	0.59	11.13
	Other	41	1.12	2.29	41	-1.95	8.84
Region	1 & 2	69	1.52	2.58	67	-0.88	10.63
	3	230	1.62	2.73	228	-0.52	10.93
	4	132	2.07	2.77	133	0.50	11.32
	Non-rated	73	1.32	2.54	71	-1.00	9.56
Community Poverty	Mid-sized	142	1.99	2.77	141	0.84	10.83
	Rural	129	1.46	2.73	130	-1.29	11.24
Community Poverty	Urban	233	1.61	2.63	228	-0.58	10.49
	Low	389	1.69	2.73	383	-0.27	10.86
High	115	1.64	2.59	116	-0.66	10.61	

Note: Group differences are statistically significant for raw gains between age groups (age 3 gained more than age 4 and 5); ethnicity groups (missing group gained greater than white and African American group), home language groups (missing group gained greater than the English group), and developmental concern groups (missing group gained greater than the non-concern group). Group differences are statistically significant for standard score gains between age groups (age 3 gained more than age 5), ethnicity groups (missing group gained greater than African American group), home language groups (missing group gained greater than the English group), and developmental concern groups (missing group gained greater than the non-concern group).

Figure C.3. WJ IV-AP standard score gains by center characteristics

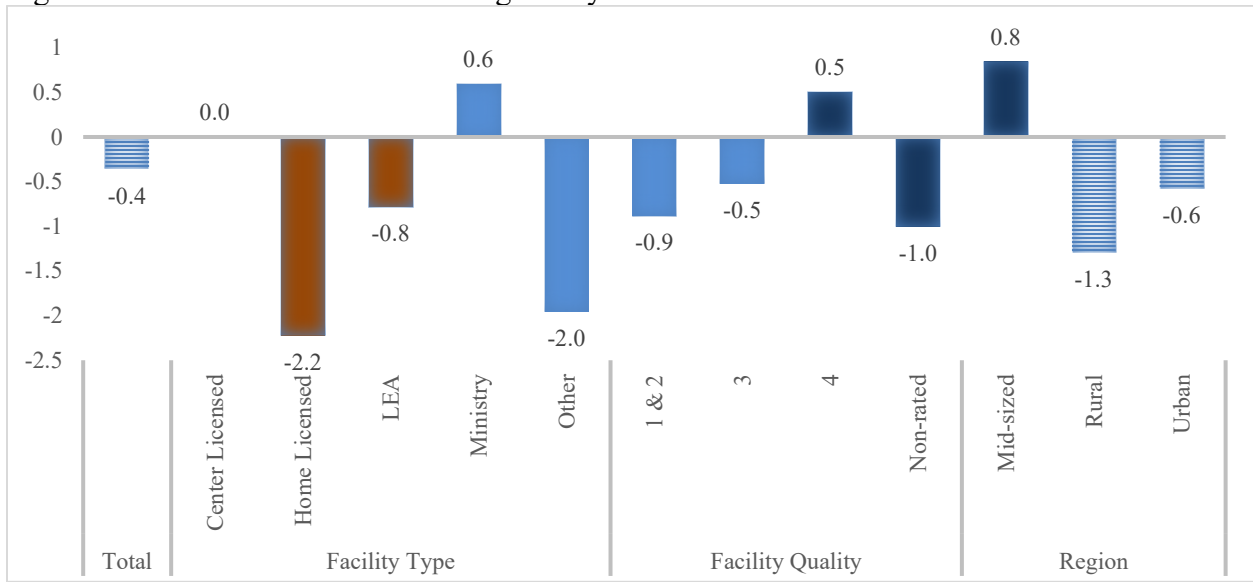


Table C.8. DCCS score gains by child and center characteristics

		DCCS Raw Score Gains			DCCS Scaled Score Gains		
		Valid N	Mean	St. Dev.	Valid N	Mean	St. Dev.
Total		306	1.40	3.97	513	0.16	0.64
Gender	Female	158	1.39	3.99	269	0.14	0.62
	Male	148	1.42	3.96	244	0.18	0.67
Age	3 and younger	83	1.76	3.89	138	0.15	0.68
	4	169	1.34	3.80	290	0.17	0.64
	5 and older	54	1.04	4.62	85	0.13	0.61
	Missing	47	1.72	3.78	58	0.21	0.64
Ethnicity	African American	19	1.74	3.49	23	0.22	0.52
	Hispanic/Latino	17	2.12	4.62	180	0.10	0.67
	Other	13	0.69	4.71	15	0.20	0.77
	White/Non-Hispanic	210	1.29	3.98	237	0.18	0.63
	Missing	4	1.50	1.73	165	0.10	0.66
Language	English	292	1.34	3.96	336	0.18	0.64
	Non-English	10	3.30	4.79	12	0.25	0.62
	Missing	2	2.00	2.83	163	0.10	0.66
Developmental Concern	No	290	1.43	3.96	336	0.19	0.64
	Yes	14	0.71	4.53	14	0.07	0.47
Facility Type	Center Licensed	93	1.73	4.47	194	0.18	0.71
	Home Licensed	65	0.89	3.44	69	0.06	0.59
	LEA	42	1.48	4.13	53	0.21	0.45
	Ministry	83	1.58	3.61	156	0.17	0.66
	Other	23	0.74	4.34	41	0.10	0.54
Facility Quality	1 & 2	53	1.09	3.85	70	0.19	0.75
	3	129	1.40	4.08	234	0.17	0.68
	4	81	1.65	4.01	136	0.18	0.53
	Non-rated	43	1.33	3.83	73	0.05	0.62
	Mid-sized	77	1.60	3.88	144	0.19	0.69
Region	Rural	75	1.55	4.33	131	0.07	0.73
	Urban	154	1.23	3.85	238	0.19	0.56
Community Poverty	Low	231	1.47	4.04	396	0.16	0.66
	High	75	1.20	3.79	117	0.15	0.58

Note: No group differences are statistically significant for raw and standard gains.

Figure C.4. DCCS score gains by center characteristics

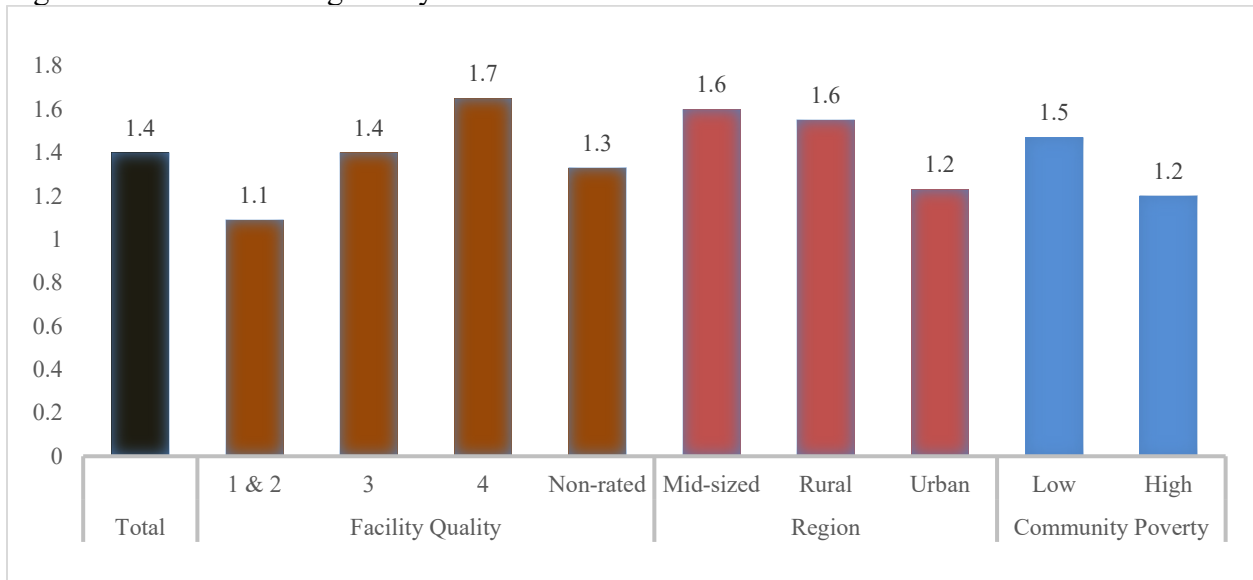


Table C.9. PT score gains by child and center characteristics

		PT Score Gains		
		Valid N	Mean	St. Dev.
Total		486	1.85	4.61
Gender	Female	255	2.27	4.54
	Male	231	1.39	4.65
Age	3 and younger	129	2.32	4.07
	4	274	2.06	4.95
	5 and older	83	0.42	3.95
	Missing	56	1.43	4.13
Ethnicity	African American	23	2.83	4.71
	Hispanic/Latino	171	2.46	4.67
	Other	15	2.87	4.79
	White/Non-Hispanic	221	1.31	4.61
	Missing	158	2.53	4.72
Language	English	316	1.41	4.51
	Non-English	12	4.58	4.03
Developmental Concern	Missing	156	2.51	4.73
	No	318	1.58	4.46
	Yes	12	0.58	6.23
Facility Type	Center Licensed	183	1.96	4.44
	Home Licensed	62	2.29	4.25
	LEA	51	1.10	4.72
	Ministry	150	1.89	4.87
	Other	40	1.50	4.87
Facility Quality	1 & 2	65	3.02	3.76
	3	222	1.27	4.89
	4	127	2.10	4.55
	Non-rated	72	2.13	4.33
Region	Mid-sized	137	1.91	4.60
	Rural	125	1.80	4.82
	Urban	224	1.84	4.51
Community Poverty	Low	377	2.04	4.57
	High	109	1.20	4.69

Note: Group differences are statistically significant for raw and standard gains between boys and girls, age groups (age 5 gained less than age 3 and 4 groups), home language groups (English group gained less than non-English and missing groups), and quality level groups (Level 1 & 2 gained more than Level 3).

Figure C.5. PT score gains by center characteristics

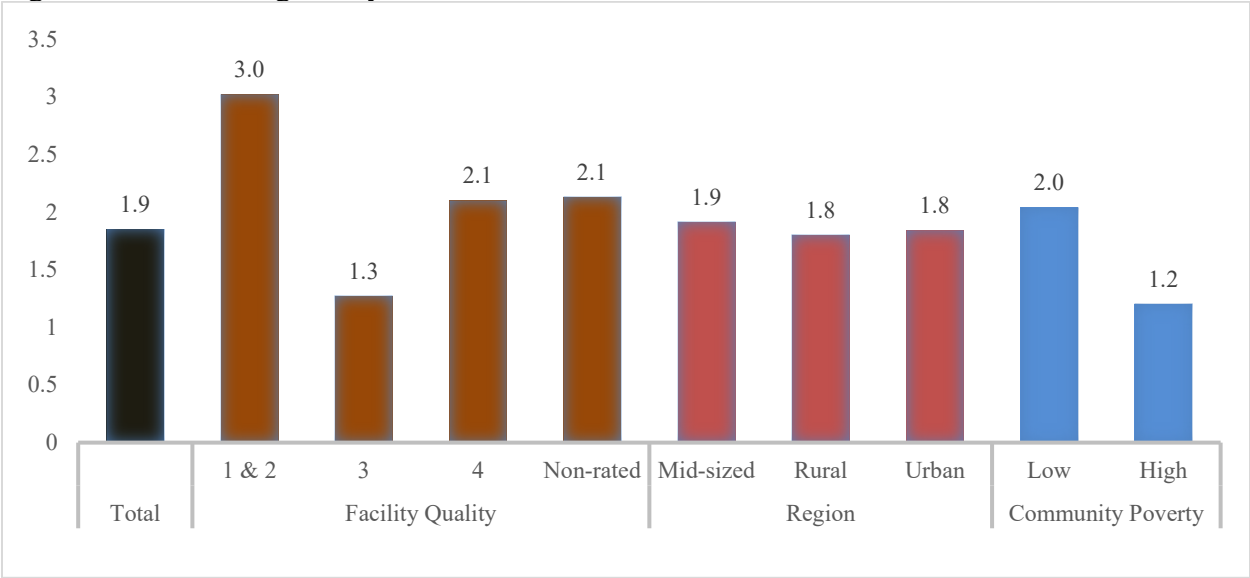


Table C.10. C-TRF Total Problems gains by child and center characteristics

		C-TRF TP Raw Score			C-TRF TP T Score		
		Valid N	Mean	St. Dev.	Valid N	Mean	St. Dev.
Total		116	-0.22	16.25	116	-1.0	8.13
Gender	Female	58	0.05	16.70	58	-1.1	7.90
	Male	58	-0.50	15.92	58	-1.0	8.43
Age	3 and younger	39	-2.41	14.02	39	-2.6	6.39
	4	57	2.56	15.79	57	0.2	8.73
	5 and older	20	-3.90	20.56	20	-1.7	9.16
	1 & 2	32	2.06	16.10	32	0.0	10.83
Facility Quality	3	56	-2.18	17.29	56	-1.6	7.19
	4	9	-1.11	12.03	9	-3.8	4.89
	Non-rated	19	2.11	15.35	19	-0.1	6.67
Region	Mid-sized	29	1.62	15.70	29	1.8	9.17
	Rural	49	-1.92	17.28	49	-2.4	6.99
Community Poverty	Urban	38	0.55	15.47	38	-1.5	8.33
	Low	98	-0.06	15.99	98	-0.9	7.80
	High	18	-1.11	18.05	18	-1.7	9.99

Note: No group differences are statistically significant for raw and standard gains.

Figure C.5. C-TRF Total Problems gains by center characteristics

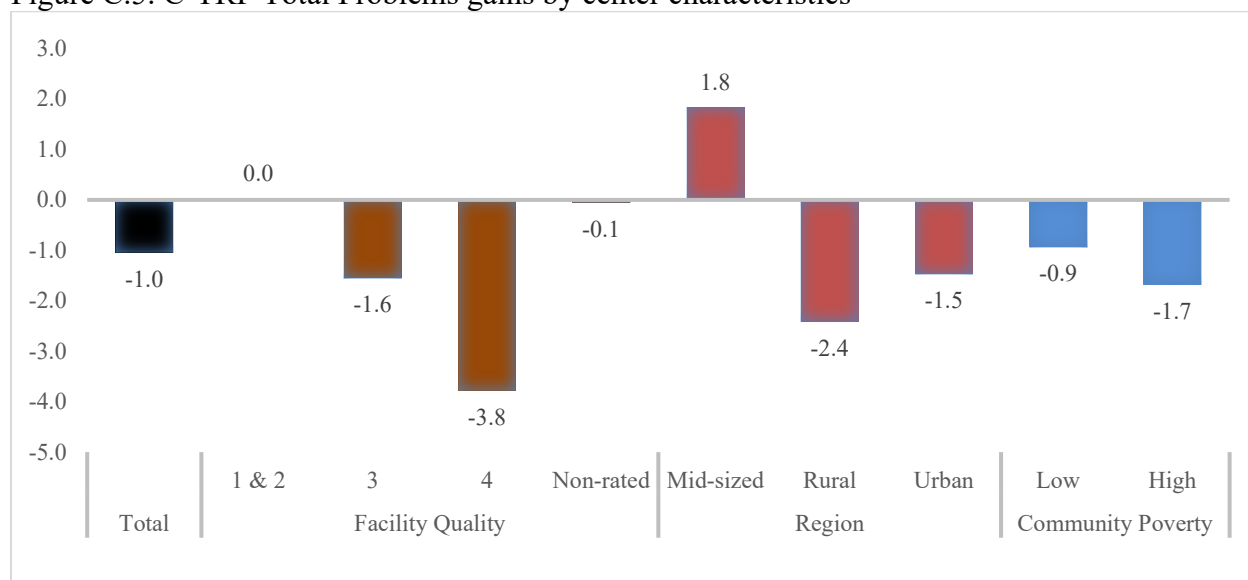




Table C.11. C-TRF Internalizing Total Problems gains by child and center characteristics

		C-TRF IP Raw Score			C-TRF IP T Score		
		Valid N	Mean	St. Dev.	Valid N	Mean	St. Dev.
Total		116	0.10	5.78	116	-0.1	8.65
Gender	Female	58	0.78	5.20	58	0.2	8.04
	Male	58	-0.57	6.27	58	-0.3	9.28
Age	3 and younger	39	0.21	5.00	39	0.0	7.21
	4	57	0.40	6.18	57	0.7	9.31
	5 and older	20	-0.95	6.15	20	-2.1	9.34
	1 & 2	32	0.41	5.38	32	0.6	10.90
Facility Quality	3	56	-0.23	6.51	56	-0.5	8.18
	4	9	0.11	3.89	9	-1.3	6.78
	Non-rated	19	0.58	5.15	19	0.7	6.75
Region	Mid-sized	29	-1.21	6.39	29	0.0	9.37
	Rural	49	0.31	5.52	49	-0.5	7.69
	Urban	38	0.84	5.59	38	0.5	9.43
Community Poverty	Low	98	-0.01	5.59	98	-0.2	8.06
	High	18	0.72	6.83	18	1.0	11.59

Note: No group differences are statistically significant for raw and standard gains.

Figure C.6. C-TRF Internalizing Total Problems gains by center characteristics

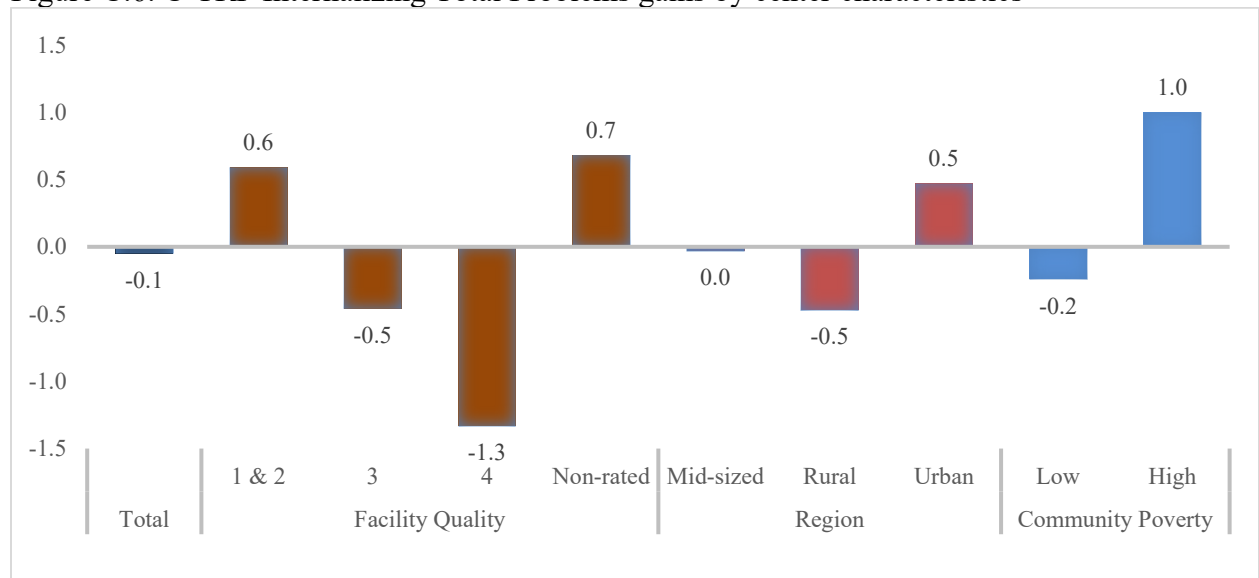
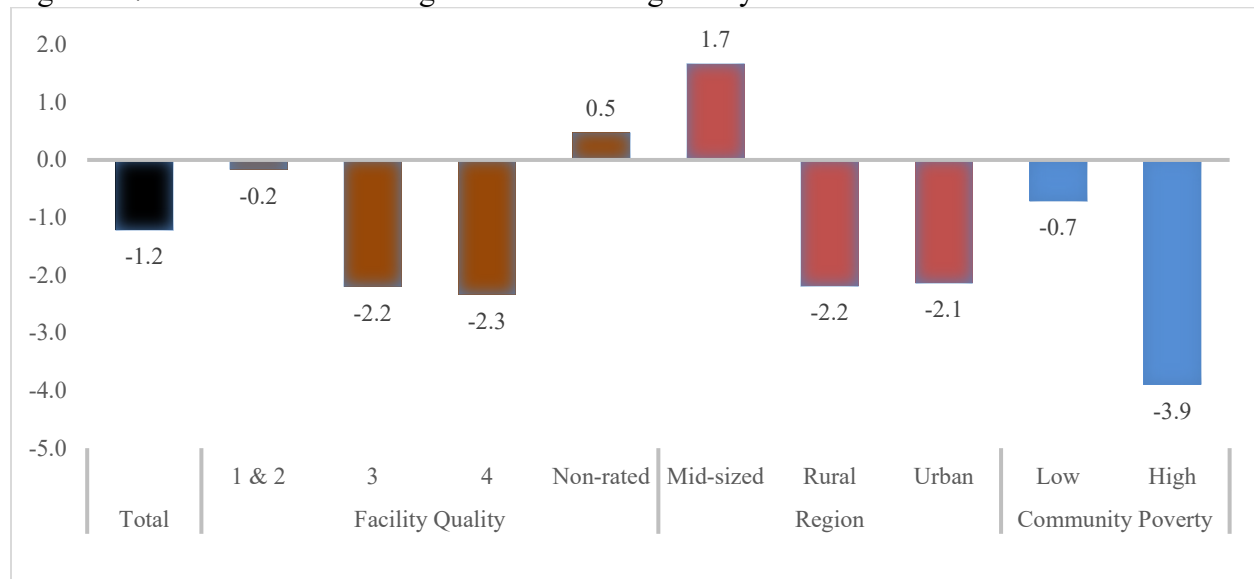


Table C.12. C-TRF Externalizing Total Problems gains by child and center characteristics

		C-TRF EP Raw Score			C-TRF EP T Score		
		Valid N	Mean	St. Dev.	Valid N	Mean	St. Dev.
Total		116	-0.52	7.77	116	-1.2	6.92
Gender	Female	58	-0.78	8.36	58	-1.0	7.20
	Male	58	-0.26	7.19	58	-1.4	6.69
Age	3 and younger	39	-2.13	6.32	39	-3.0	5.40
	4	57	1.16	7.34	57	0.1	7.55
	5 and older	20	-2.15	10.5	20	-1.3	7.22
	1 & 2	32	0.81	6.99	32	-0.2	7.70
Facility Quality	3	56	-1.79	8.47	56	-2.2	6.58
	4	9	-0.33	7.52	9	-2.3	7.00
Region	Non-rated	19	0.89	6.85	19	0.5	6.40
	Mid-sized	29	1.93	7.06	29	1.7	7.47
	Rural	49	-1.65	8.41	49	-2.2	6.36
Community Poverty	Urban	38	-0.92	7.17	38	-2.1	6.75
	Low	98	-0.17	7.78	98	-0.7	6.74
	High	18	-2.39	7.63	18	-3.9	7.48

Note: Group differences are statistically significant for standard gains between different regions (Mid-sized region gained greater than rural and urban areas).

Figure C.7. C-TRF Externalizing Total Problems gains by center characteristics



## Appendix D. Parent surveys

Table D1. Parent characteristics

Parent Characteristics (n=324)	Total	Percent
<b>Region</b>		
Mid-sized	92	28.4%
Rural	104	32.1%
Urban	128	39.5%
<b>Auspice</b>		
Center Licensed	96	29.6%
Home Licensed	72	22.2%
LEA	31	9.6%
Ministry	112	34.6%
Other	13	4.0%
<b>Rating</b>		
Non-Rated	43	13.3%
1 & 2	50	15.4%
3	138	42.6%
4	93	28.7%
<b>Child's Race/Ethnicity</b>		
Asian	2	0.6%
Black, African, or African American	29	9.0%
Hispanic/Latino	9	2.8%
White	257	79.6%
Multi-Racial	15	4.6%
Other	11	3.4%
<b>Child Age Range</b>		
Infant/Toddler	97	29.9%
Pre-K	227	70.1%
<b>Household Income</b>		
Less than \$30,000	42	13.2%
\$30,001 to \$50,000	52	16.4%
\$50,001 to \$100,000	93	29.3%
\$100,001 to \$150,000	64	19.8%
\$150,000+	66	20.8%
<b>Total</b>	<b>324</b>	

Table D2. Parent characteristics

Parent Characteristics (n=324)	Total	Percent
<b>Highest Level of Education Completed</b>		
Did Not Finish High School	7	2.2%
High School Diploma/GED	26	8.0%
Some College	64	19.8%
Associate's Degree	37	11.5%
Bachelor's Degree	104	32.2%
Master's Degree or Higher	85	26.3%
<b>Age</b>		
18-20	1	0.3%
21-25	13	4.0%
26-29	46	14.2%
30-34	114	35.2%
35-39	102	31.5%
40-44	33	10.2%
45-49	7	2.2%
50-54	3	0.9%
60+	5	1.5%
<b>Total</b>	<b>324</b>	

Table D3. Parental Perceptions of Teacher Practices by PTQ

		Talks to me at least once a week.	Teaches my child how to get along with others.	Teaches my child to control emotions and behaviors	Teaches my child academic skills	Tracks my child's progress	Is fluent in my child's primary home language
<b>Don't Know</b>	0/Non-Rated	0%	9.30%	14.00%	0.00%	7.00%	0.00%
	1 & 2	0%%	6.00%	8.00%	2.00%	18.00%	0.00%
	3	0%	4.40%	5.10%	2.90%	8.00%	0.00%
	4	0%	4.30%	9.80%	2.20%	5.40%	1.10%
<b>Not at All</b>	0/Non-Rated	0%	0.00%	0.00%	0.00%	2.30%	0.00%
	1 & 2	4%	0.00%	0.00%	2.00%	2.00%	0.00%
	3	1.40%	0.00%	0.70%	1.40%	0.00%	0.00%
	4	1.10%	0.00%	0.00%	0.00%	0.00%	0.00%
<b>Very Little</b>	0/Non-Rated	14%	0.00%	0.00%	0.00%	2.30%	0.00%
	1 & 2	4%	0.00%	2.00%	2.00%	4.00%	0.00%
	3	2.90%	0.00%	1.40%	0.70%	0.70%	0.70%
	4	4.40%	1.10%	1.10%	0.00%	0.00%	0.00%
<b>Somewhat</b>	0/Non-Rated	11.60%	11.60%	20.90%	14.00%	30.20%	2.30%
	1 & 2	8%	6.00%	14.00%	26.00%	12.00%	0.00%
	3	13.80%	5.90%	10.10%	8.70%	14.50%	1.50%
	4	13.20%	4.30%	5.40%	13.00%	9.80%	3.30%
<b>A Lot</b>	0/Non-Rated	27.90%	37.20%	39.50%	34.90%	23.30%	4.70%
	1 & 2	14%	28.00%	32.00%	26.00%	22.00%	14.00%
	3	17.40%	21.30%	22.50%	24.60%	29.70%	12.40%
	4	14.30%	32.60%	32.60%	21.70%	25.00%	12.00%
<b>Very Much</b>	0/Non-Rated	46.50%	41.90%	25.60%	51.20%	34.90%	93.00%
	1 & 2	70%	60.00%	44.00%	42.00%	42.00%	86.00%
	3	64.50%	68.40%	60.10%	61.60%	47.10%	85.40%
	4	65.60%	57.60%	51.10%	63.00%	59.80%	83.70%

Table D4. Parental Perceptions of Teacher Practices by Child Age

		Talks to me at least once a week.	Teaches my child how to get along with others.	Teaches my child to control emotions and behaviors.	Teaches my child academic skills.	Tracks my child's progress.	Is fluent in my child's primary home language.
<b>Don't Know</b>	Infant/Toddler	0%	6.3%	13.4%	5.2%	14.4%	0.00%
	Pre-K	0%	4.9%	5.8%	0.9%	6.2%	0.4%
	<b>Total</b>	<b>0%</b>	<b>5.3%</b>	<b>8.0%</b>	<b>2.2%</b>	<b>8.7%</b>	<b>0.3%</b>
<b>Not at All</b>	Infant/Toddler	0%	0.0%	1.0%	1.0%	2.1%	0.0%
	Pre-K	2.2%	0.0%	0.4%	0.9%	0.0%	0.0%
	<b>Total</b>	<b>1.6%</b>	<b>0.0%</b>	<b>0.3%</b>	<b>0.9%</b>	<b>0.6%</b>	<b>0.0%</b>
<b>Very Little</b>	Infant/Toddler	5.2%	0.0%	1.0%	1.0%	0.0%	0.0%
	Pre-K	4.9%	0.04%	1.3%	1.0%	1.8%	0.4%
	<b>Total</b>	<b>5.0%</b>	<b>0.03%</b>	<b>1.2%</b>	<b>0.6%</b>	<b>1.2%</b>	<b>0.3%</b>
<b>Somewhat</b>	Infant/Toddler	7.2%	11.5%	15.5%	21.6%	19.6%	4.2%
	Pre-K	14.7%	4.0%	8.8%	9.7%	12.8%	0.9%
	<b>Total</b>	<b>12.4%</b>	<b>6.2%</b>	<b>10.8%</b>	<b>13.3%</b>	<b>14.9%</b>	<b>1.9%</b>
<b>A Lot</b>	Infant/Toddler	16.5%	27.1%	26.8%	25.8%	23.7%	7.3%
	Pre-K	17.8%	28.0%	30.1%	25.2%	27.4%	13.3%
	<b>Total</b>	<b>17.4%</b>	<b>27.7%</b>	<b>29.1%</b>	<b>25.4%</b>	<b>26.3%</b>	<b>11.5%</b>
<b>Very Much</b>	Infant/Toddler	71.1%	55.2%	43.3%	45.4%	40.2%	88.5%
	Pre-K	60.4%	62.7%	53.5%	62.8%	51.8%	85.0%
	<b>Total</b>	<b>63.7%</b>	<b>60.4%</b>	<b>50.5%</b>	<b>57.6%</b>	<b>48.3%</b>	<b>86.0%</b>