NIEER
NATIONAL INSTITUTE FOR EARLY EDUCATION RESEARCH

EVALUATION OF THE PHILADELPHIA PREK PROGRAM

Classroom Quality Report

September 2017

Milagros Nores, PhD, Jessica Francis, PhD & W. Steven Barnett, PhD.

Table of Contents

Study Background

This report is presented as the first of the evaluation study of Philadelphia's PreK Program (PHLpreK), conducted by the National Institute of Early Education Research (NIEER). The evaluation study examines several key research questions on the effectiveness of PHLpreK on impacting children's learning outcomes and ultimately the overall economy, and encompasses how those outcomes relate to classroom quality. This report provides the City and PHMC a detailed account of the classroom quality for students in PHLpreK classrooms and provides a comprehensive understanding of the environment and teaching practices in the classrooms. This level of detail provides the opportunity for PHMC to use a data-driven continuous improvement approach to support increased quality.

Introduction

High-quality preschool education has been put forward as a response to one of the nations most persistent educational problems: the achievement gaps observed in children from minority and low-income families (Ceci & Papierno, 2005; Duncan & Murnane, 2011; Nores & Barnett, 2015). Research has shown for some time now that that high-quality preschool education programs can produce lasting effects in school success and achievement (Barnett, 2008; Camilli, Vargas, Ryan & Barnett, 2010; Yoshikawa et al., 2013). Well-designed preschool education programs have produced gains strong enough to close half the achievement gap between children from low- and high-income families at kindergarten entry and stronger reductions in gaps for minorities (Camilli et al., 2010, Friedman-Kraus, et al., 2016). Some rigorous studies have also shown that long-term impacts on achievement and social development can persist even if these may be somewhat smaller than short-term impacts (Barnett, 2011; Camilli, et al., 2010).

As a consequence, how to define quality and to what extent preschool programs' structural characteristics and process indicators influence program effectiveness on children has become a central policy aspect to early childhood systems. Weak associations between structural features of preschool programs and children's learning have led researchers to focus on classroom process and in-service professional development to improve effectiveness (Pianta & Hamre, 2009; Hamre, et al, 2014). The field has grown in the utilization of observational measures of quality as part of continuous improvement cycles, quality rating systems, and for program evaluation (Martinez-Beck, 2011). Observational measures have also become central in Head Start evaluations (U.S. Department of Health and Human Services, Administration for Children and Families, 2010). While the strength of association of child outcomes from existing measures of quality has been indicated to be modest (Burchinal, Kainz & Cai, 2011), the associations have been shown to be stronger at higher levels of process quality (Burchinal, Kainz & Cai, 2011; Burchinal, et al., 2014; Hatfield, Burchinal, Pianta & Sideris, 2016; Hatfield, et al., 2015 Hestenes, et al., 2015; Weiland & Yoshikawa, 2013). Hatfield, et al. (2015) concludes that stronger findings between quality and language and literacy skills to be in classrooms language-rich classrooms. There is also some evidence on positive associations with children's behavioral skills and executive functions (Weiland, Ulvestad, Sachs & Yoshikawa, 2013) or socio-emotional functioning (Hestenes, et al., 2015).

Philadelphia PreK Program

Philadelphia is a more diverse, safe, educated, and thriving city than it has been in decades. However, its potential for future progress will be held back until all children, regardless of race, ethnic or socioeconomic backgrounds are ready to succeed in school. Quality pre-kindergarten is one way to make this happen. In the long term, quality pre-K programs result in fewer children in special education classes, higher graduation rates, increased earning potential, better health, and narrowing the achievement gap rooted in income and other environmental factors. Quality pre-K programs also have been shown to support parents' efforts to enter into and thrive in the workforce. The benefits ripple through the workforce, economy, and community. Yet only one in three of Philadelphia's 42,500 three- and four-year-olds have access to affordable quality pre-K (Philadelphia Commission on Universal Pre-Kindergarten, 2016).

On May 19, 2015, Philadelphians voted overwhelmingly (80 percent) to create the Philadelphia Commission on Universal Pre-Kindergarten. Its charge: propose a universal pre-K program that provides quality, affordable, and accessible services to three- and four-year-olds throughout Philadelphia. The Commission was tasked with providing recommendations for sustainable funding for pre-K *that does not reduce existing K–12 funding.*

In order to produce and sustain effective, high-quality pre-K programs, the commission has recommended that the program is evaluated by an external entity to monitor quality and report on child outcomes. This recommendation was based on a review of scientific research, evidence-based practices, and current local and state efforts. Additionally, the commission garnered input from experts in the field, members of the public, through multiple hearings, community meetings, discussions, and surveys. The National Institute for Early Education Research (NIEER) serves as an external evaluator and is conducting a multi-year, multi-site study that employs a combination of methods and designs to assess the program components, program quality, and impacts on children's learning and development. The current report is focused on classroom quality.

Results showed that PHLpreK classrooms are averaging moderate levels of quality in areas of emotional support and classroom organization for children. In contrast, there is work to be done to support teachers in their use of strategies and techniques for scaffolding and expanding children's learning and language development. Observations also revealed that classrooms are generally effective at implementing a variety of activity settings. Children spend over a third of their day with various opportunities for learning. However, about a fifth of the day is spent in transitions, where there are few opportunities for learning. Children also engage in different content areas for some portion of the day, but receive no learning content for a large portion of the day. Classrooms exhibit a balance of didactic and scaffolded interactions, though children are rarely asked to explain or justify their thinking through metacognitive processes. Quality scores were examined separately for several subgroups of interest, including star level, number of PHL contracted slots, number of PHLpreK classrooms, lead teacher credential, PHLpreK partner agency, council district, and curriculum. Minimal differences are found between subgroups. Higher quality classrooms arethose that spend less time in transition, more time with content areas, and integrated content more often.

Study Methods

Classroom quality data was collected in the last four months of the school year, between March through June of 2017. The purpose was to assess quality of the classrooms through using two different observation tools, during two classroom visits of approximately 2.5–3 hours each. This was complemented with a checklist developed by NIEER to code the presence (or lack) of materials in the classroom.

1. Sample

The study includes all PHLpreK classrooms in the city. Generalization is dependent on the similarities among programs in the city. The sample size for the study was 139 pre-K classrooms, 13 of which were housed in home-based programs. CLASS data were collected in all 139 classrooms (both center-based and home-based) in two separate visits. EduSnap data were collected in all 126 center-based classrooms.

2. Measures

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

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

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

Table 1. CLASS Domains and Dimension Descriptions.

Domain	Dimension	Description
Emotional Support	Positive Climate	Reflects the emotional connection between teachers and children and among children, and the warmth, respect, and enjoyment communicated by verbal and nonverbal interactions.
	Negative Climate	Reflects the overall level of expressed negativity in the classroom: frequency, quality, and intensity of teacher and peer negativity.
	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 teacher's interactions with students and classroom activities place an emphasis on students' interests, motivations, and points of view and encourage student responsibility and autonomy.
Classroom Organization	Behavior Management	Encompasses the teacher's ability to provide clear behavior expectations and use effective methods to prevent and redirect misbehavior.

	Productivity	Considers how well the teacher manages instructional time and routines and provides activities for students so that they have the opportunity to be involved in learning activities.
	Instructional Learning Formats	Focuses on the ways in which teachers maximize students' interest, engagement, and abilities to learn from lessons and activities.
Instructional Support	Concept Development	Measures the teacher's use of instructional discussions and activities to promote students' higher-order thinking skills and cognition with a focus on understanding rather than 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.

EduSnap Classroom Observation (Ritchie, Weiser, Mason, & Holland, 2015)

EduSnap is a tool that provides quantification of the experiences that children have throughout the school day. This measure provides an in-depth look at how students experience their day by documenting the actual time students spend in activity settings (e.g., whole group, free choice, transitions), content areas (e.g., reading, science, math), student learning approaches (e.g., collaboration, meta cognition), and teaching approaches (e.g., didactic, scaffolds). Data provide insight into curriculum balance, curriculum integration, and interactions between teachers and children. Observers coded classrooms on EduSnap for three hours in each classroom, starting first thing in the morning.

High-quality classrooms exhibit a balance across activity settings, content areas, and student learning and teaching approaches to best provide all children with a variety of experiences across the school day and within scheduled time blocks. The following information related to EduSnap is meant to help with the interpretation of observation results provided in this report.

- Access to a variety of Activity Settings is important as children optimize their learning in different ways—some have greater or less tolerance for large and small groupings, noise level, auditory, visual, tactile and kinesthetic input.
- Frequency of exposure to learning opportunities increases children's academic achievement. Developing literacy and math skills and processes is essential for success at all levels of learning and should be emphasized according to children's needs and developmental levels. However, these should be balanced with the importance of learning foundational knowledge in all subject areas, including science, social studies, art, and music. A well integrated curriculum would allow access to multiple content areas to children and would strengthen learning across and within them.
- Providing children many opportunities to work together and to engage in metacognitive thinking supports both their social/emotional and academic development.
- Didactic instruction provides children with needed practice and repetition that helps them build their skill base across the curriculum, giving them models, demonstrations, information and guidance.
- Scaffolded instruction involves teachers asking open-ended questions, engaging in feedback loops, and probing more deeply into children's thinking and understanding. This type of instruction enables the teachers to know specifically how much the children understood from a lesson, identify and remediate group or individual misunderstandings, and engage children in the learning process. Knowing this enables the teacher to respond

by modifying the current and subsequent learning experiences and activities according to the individual and group needs. Both didactic and scaffolded instruction are important teaching styles and should be incorporated in a balanced fashion throughout the course of each school day and within each lesson.

Classroom Observation Checklist

The Classroom Observation Checklist is an abbreviated version of the ECERS-R that assesses some of the environmental features that are present in the classroom. Each item is recorded as 'yes' or 'no' and no actual scores are assigned. The measure is broken down into five main areas including furniture, room arrangement, classroom display, health and safety, and daily schedule. The classroom inventory of materials across different areas of learning is also recorded. This instrument is displayed in the appendix (Table I).

3. Procedures

Trained and reliable observers conducted the observations of classroom quality. Training was provided in administering the observation protocol that includes EduSnap and the CLASS for all center-based preschool classrooms. Home-based classrooms were observed with the CLASS only. Training took place in two separate two-day workshops. EduSnap observers were trained by the developer of the EduSnap and were subsequently required to complete kappa reliability with pre-coded videos online. CLASS observers were trained by a CLASS certified trainer and met the Teachstone reliability requirements (80%) for observer certification. All observation score sheets were cleaned and entered at NIEER by trained staff.

Results

We present the results by instruments, first for the CLASS, then for EduSnap. In addition, information for the Checklist was delivered to PHMC separately and is not included in this report.

1. CLASS

Table 2 below presents aggregate results across all PHLpreK classrooms for each CLASS dimension and domain. The scoring patterns with instructional support scoring lower than other domains in consistent with the field and other preschool programs. Pre-K CLASS mean scores were 5.85 for Emotional Supports (ES), 5.34 for Classroom Organization (CO) and 2.41 for Instructional Supports (IS). A discussion of each domain is subsequently presented. Some research appears to support (Burchinal et al. 2009; OPRE, 2010) thresholds for ES and CO above 5 and IS above 3 as necessary for a relation between quality and children's outcomes to be evidenced. In PHLpreK, 85% of the classrooms were above these thresholds in ES, 71% were above them in CO, and only 19% were above the threshold of 3 in IS.

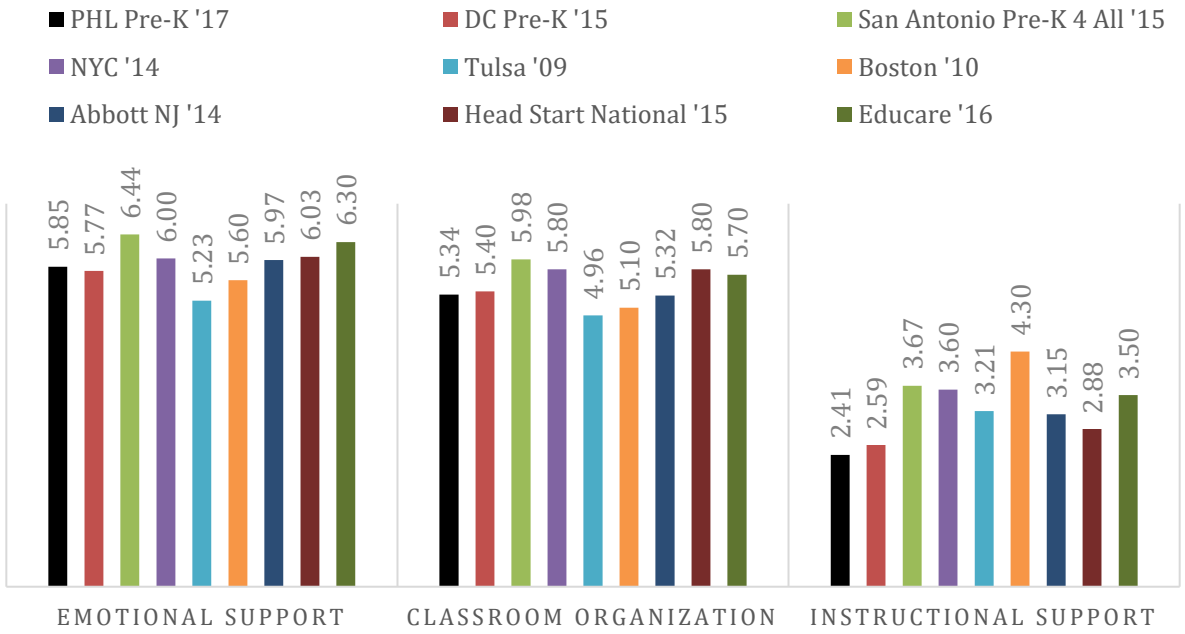
Table 2. PreK CLASS Dimension and Domain Means and Ranges, N = 139.

CLASS Dimensions and Domains	Mean	Minimum	Maximum
Emotional Support Domain (ES)	5.85	2.85	6.90
1. Positive Climate	5.90	1.60	7.00
2. Negative Climate*	6.77	5.00	7.00
3. Teacher Sensitivity	5.69	2.20	7.00
4. Regard for Student Perspectives	5.03	2.00	6.80
Classroom Organization Domain (CO)	5.34	1.87	6.93
5. Behavior Management	5.49	1.60	7.00
6. Productivity	5.76	1.80	7.00
7. Instructional Learning Formats	4.77	1.60	7.00
Instructional Support Domain (IS)	2.41	1.00	5.00
8. Concept Development	2.09	1.00	4.80
9. Quality of Feedback	2.23	1.00	5.00
10. Language Modeling	2.91	1.00	5.20

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

The score patterns for the PHLpreK follow those of the National Overview of CLASS in pre-K classrooms of 2015 (OHS, 2015), with the highest scores being prevalent in the emotional support domain (national mean of 6.03), followed by the Classroom Organization domain (national mean of 5.80) and with the lowest scores being in the Instructional Support domain (national mean of 2.88). The PHLpreK CLASS scores together with those from various other programs in the U.S. are shown in Figure 1 for comparison purposes, including high-quality city programs.

Figure 1. Comparison of PHLpreK CLASS scores with other programs.

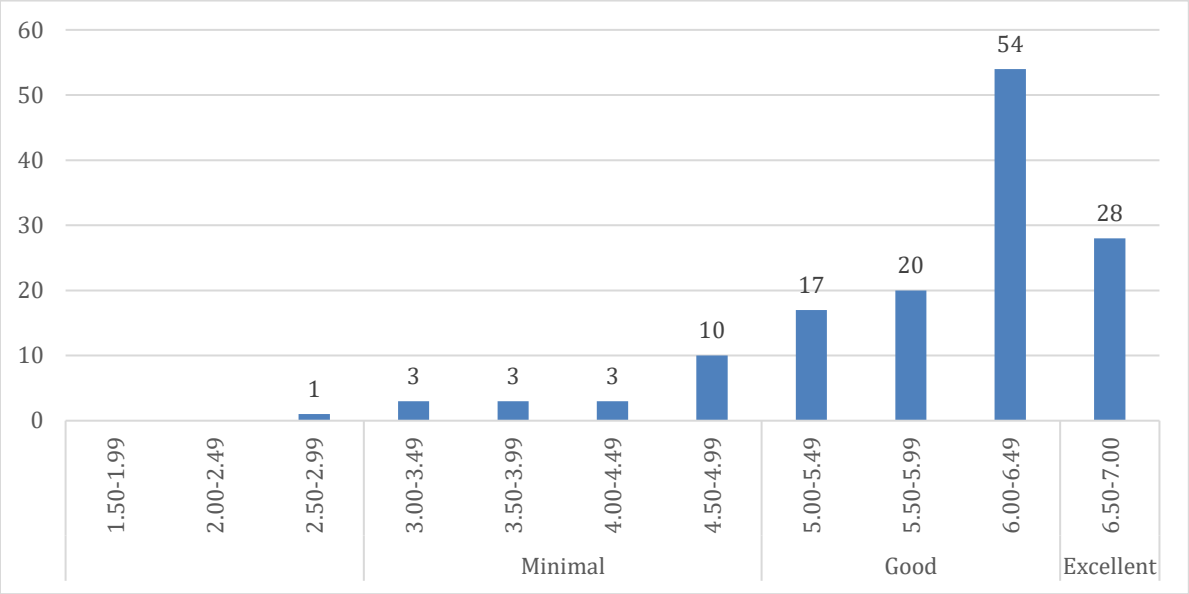


Emotional Support Domain

The overall mean score for ES is 5.85 (SD 0.82), putting it in the high end of the mid range. The minimum score is 2.85, which indicates there was at least one classroom in which there was a low level of emotional support throughout all five cycles observed. The highest scoring dimension is Negative Climate, with a mean of 6.77, indicating most classrooms exhibited very little negative interaction between teachers and children and between children. The lowest scoring dimension is Regard for Student Perspectives, with a mean of 5.03. A mid-range score in this dimension indicates classrooms with teachers who sometimes show flexibility, sometimes give students responsibility, but sometimes are restrictive of students' movement throughout the day.

Figure 2 presents the distribution of scores across ES. Out of the 139 classrooms observed, 19 classrooms scored somewhere in the mid range (3.00-4.99, 14%) and 119 classrooms scored somewhere in the high range (5.00-7.00, 86%).

Figure 2. Distribution of Scores for Emotional Support Domain, N = 139

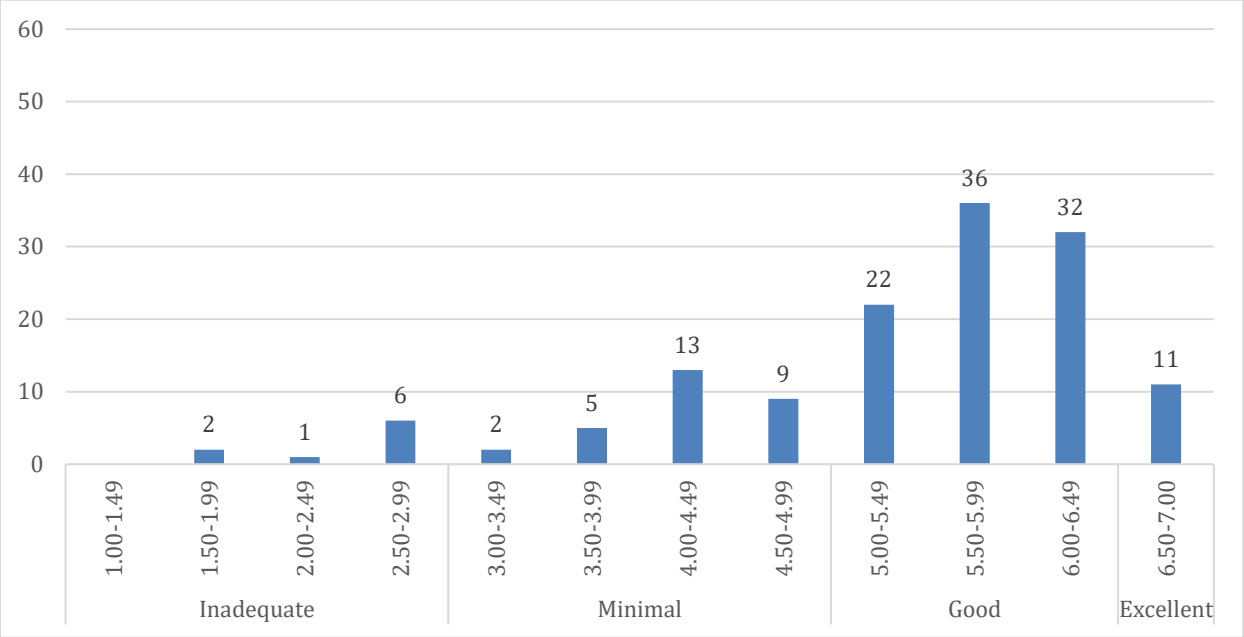


Classroom Organization Domain

The overall mean score for the Classroom Organization Domain is 5.34 (SD 1.11). Behavior Management and Productivity are the two higher scoring dimensions, with means of 5.47 and 5.73, respectively. Such high scores indicate classrooms in which there are effective methods in place to both prevent and redirect misbehavior, while most student behavior observed during the five cycles was compliant and appropriate. Additionally, the teachers were observed to manage their instructional time well, with little time wasted. It is important to note that in the Productivity dimension, the quality of the activities is not considered, rather only that there are activities available.

Figure 3 presents the distribution for the mean scores in the Classroom Organization Domain. Like the Emotional Supports Domain, very few classrooms scored in the low range in all three dimensions. Of the 139 classrooms observed, 38 of them scored in the mid range (27%), and 10 classrooms scored in the high range above 5 (73%).

Figure 3. Distribution of Scores for Classroom Organization Domain, N = 139.

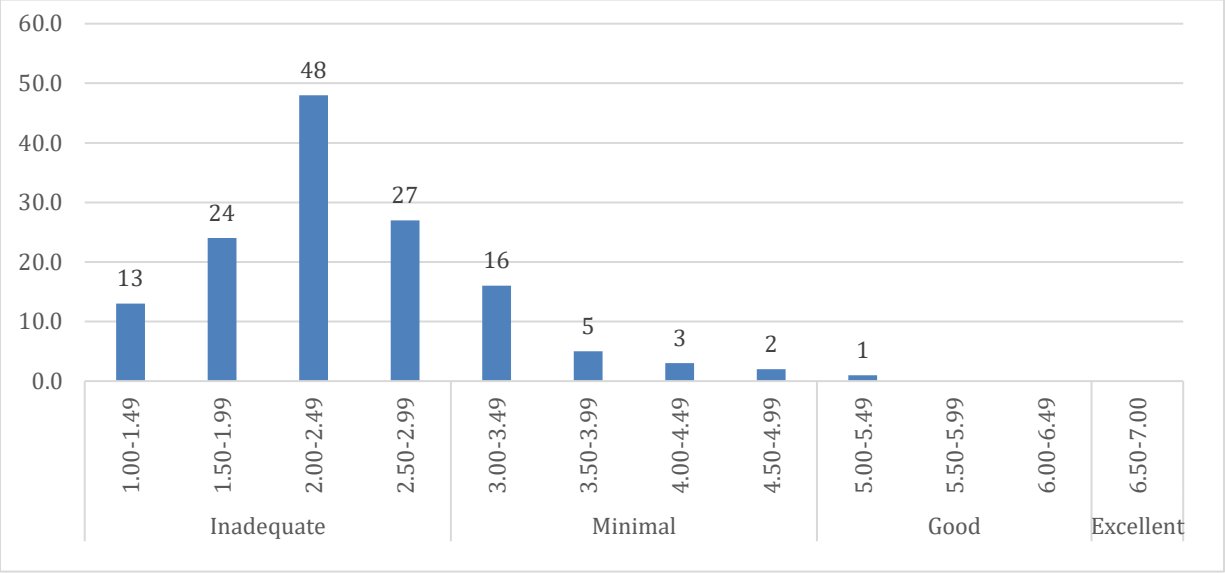


Instructional Support Domain

IS assesses the interactions through which teachers deliver and facilitate high-order thinking skills, and develop language. As mentioned previously, this domain is the most difficult, yet critically important, when considering teacher practices that have impacts on student growth. The mean score for this domain is 2.41 (SD 0.75) with averages ranging from 1 to 5 on a 7-point scale.

Figure 4 presents the distribution for the mean scores in the Instructional Support Domain. Classrooms cluster near the lower end of the scale for this item. The standards for this domain for inadequate versus minimal are the same as those for the other two dimensions. Having said this, quality in IS tends to be lower than for the other two dimensions for all programs and the thresholds found in the literature are lower. To set a standard of what is higher versus lower quality for IS, we have set these cutoffs around 3 to help interpret these results. There are 27 classrooms/home-providers scoring above 3 for this domain (20%), and 112 classrooms scoring under 3 (80%).

Figure 4. Distribution of Scores for Instructional Support Domain, N = 139.



CLASS results are also presented for the subgroups of interest in Table 3. The patterns follow those of the overall sample, with ES scores being higher, followed by CO scores, and with IS scores trailing quite behind. Patterns that emerge more clearly are that centers with a STAR rating of 4 score higher across all three domains, and it is particularly noticeable in IS, especially when compared to centers with a STAR rating of 1.

Table 3. CLASS domain scores by subgroups, N = 139.

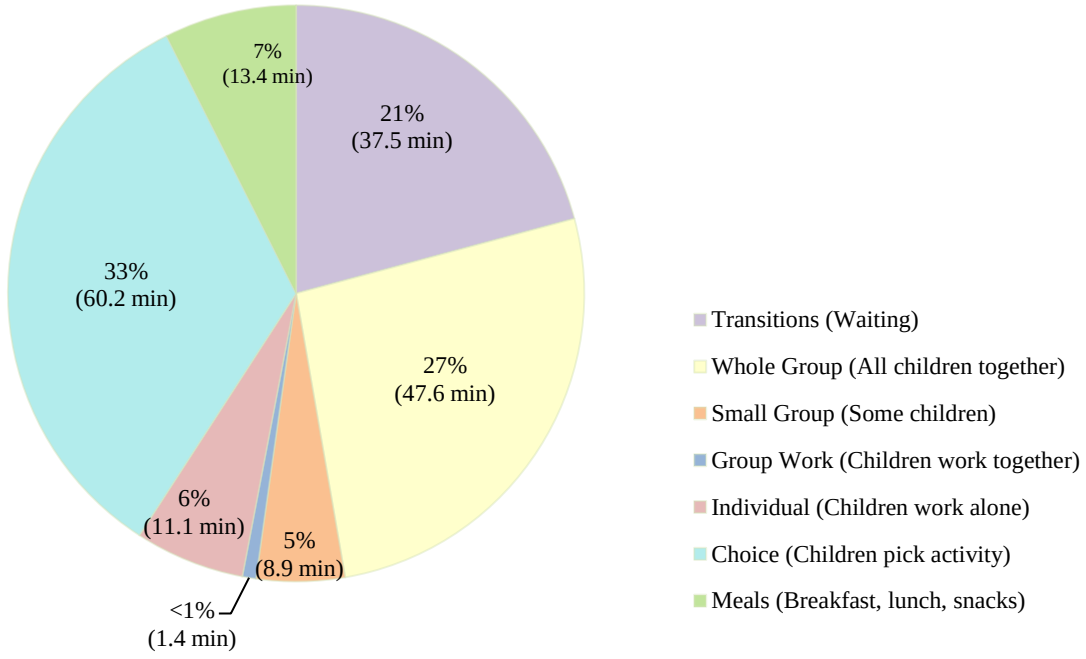
		Mean Scores		
		Emotional Support	Classroom Organization	Instructional Support
STAR Level	1 (n=20)	5.79	5.34	2.29
	2 (n=25)	5.76	5.12	2.41
	3 (n=58)	5.79	5.33	2.30
	4 (n=29)	6.00	5.45	2.55
	Unrated (n=7)	6.17	5.78	3.07
Number of PHL Contracted Slots	0-18 (n=37)	5.88	5.38	2.69
	19-28 (n=41)	5.73	5.26	2.29
	29-48 (n=28)	5.94	5.41	2.32
	49-90 (n=33)	5.88	5.33	2.33
Number of PHL preK classrooms	1 (n=68)	5.73	5.30	2.45
	2 (n=30)	5.90	5.33	2.47
	3 (n=21)	6.11	5.52	2.44
	4 or more (n=20)	5.88	5.34	2.16
Lead Teacher Credential	CDA/ECE Course (n=4)	6.16	6.17	2.85
	AA (n=32)	5.83	5.34	2.27
	BA (n=45)	5.87	5.38	2.36
	MA (n=2)	6.63	6.03	3.33
PHLpreK Partner Agency	District 1199C (n=13)	5.96	5.56	2.67
	Phila SD (n=14)	5.73	5.55	2.24
	PHMC (n=51)	5.91	5.41	2.45
	UAC (n=61)	5.79	5.19	2.36
Council District	1 (n=5)	4.88	3.80	2.41
	2 (n=26)	6.02	5.41	2.28
	3 (n=25)	5.81	5.43	2.33
	4 (n=3)	6.03	5.89	2.42
	5 (n=9)	5.94	5.59	2.31
	6 (n=18)	5.99	5.66	2.63
	7 (n=15)	5.84	5.35	2.54
	8 (n=18)	5.88	5.25	2.40
	9 (n=12)	5.63	5.09	2.27
	10 (n=8)	5.74	5.18	2.69
Curriculum	Creative (n=75)	5.88	5.34	2.36
	Creative + another (n=13)	5.95	5.67	2.53
	Mother Goose Time	5.79	5.34	2.48
	Other (n=25)	5.79	5.15	2.44

2. EduSnap

EduSnap results are presented for the aggregate sample and also separately for the following subgroups: STAR rating level, number of PHL contracted slots, number of classrooms per site, teacher credentials, PHLpreK partner agency, council district, and curriculum.

A first way of looking at children's experiences of the day is thinking how children experience activity settings, which is defined as the groupings that occur on a minute-by-minute basis (Figure 5). These groupings include whole group, small group, transitions, group work, meals, and individual and choice time. The time spent in choice (32%) and whole group (27%) are on average well balanced, while the time spent on transitions (21%) is high and the time spent in small groups (5%) is quite low. Group work activities are geared more toward children in later grades so these percentages for this pre-K sample are of less importance. Converting time to minutes is helpful in interpreting how these percentages translate throughout the day. One percent is on average equivalent to 1.8 minutes, five percent to 9 minutes, ten percent to 18 minutes, and so forth. Figure 5 also includes minutes following this scale.

Figure 5. Percentages of Time and Minutes Spent in Activity Settings, n=139.



Patterns across subgroupings of programs, whether by STAR ratings, number of contracted slots, curriculum, or number of PHLpreK classrooms, do not differ much (Table 4). Transitions generally are around 20% and even up to about 25% in some programs without any clear patterns. Whole group activities vary between 20 and 37% while small group activities are generally low across the board (around 5%). Choice is around 30-35% for most types of centers, and slightly lower in STAR rating 1 centers and centers where PHLpreK has contracted over 76 slots (but this only includes 2 centers). Additionally, STAR rating 1 centers seem to rely much more on whole group than most other types of centers. Transition times are shorter in classrooms

with teachers with a CDA/ECE or an MA/ME (although information on teacher credentials was only available for 60% of teachers in the sample, so comparisons across teacher credentials in this report need to be interpreted with caution).

Table 4. Percentages of time spent in Activity Settings by Subgroups, n=126.

		% of Time						
		Meals	Transitions	Whole Group	Small Group	Individual	Choice	Group Work
STAR Ratings	1 (n=16)	4.64	21.41	31.80	4.67	5.61	31.66	0.22
	2 (n=21)	7.65	22.82	19.45	6.38	8.80	34.05	0.87
	3 (n=58)	7.80	20.26	26.71	4.59	5.81	33.80	1.03
	4 (n=25)	8.06	20.82	27.92	3.77	4.94	33.93	0.56
	Unrated (n=6)	7.84	17.62	27.18	9.77	7.20	30.29	0.11
Number of PHL Contracted Slots	0-18 (n=25)	8.92	20.41	24.74	6.71	5.17	33.75	0.31
	19-28 (n=38)	7.78	20.03	27.49	5.58	5.23	33.28	0.61
	29-48 (n=30)	5.70	21.14	26.32	3.96	6.01	36.25	0.62
	49-90 (n=33)	7.47	21.75	26.67	3.84	8.10	30.75	1.42
Number of PHL preK classrooms	1 (n=40)	8.36	19.37	27.82	6.34	5.65	31.84	0.61
	2 (n=40)	5.87	20.94	24.99	5.41	5.55	36.25	0.99
	3 (n=24)	7.35	22.84	26.73	3.19	3.33	35.68	0.88
	4 or more (n=22)	8.57	21.08	26.26	3.49	11.36	28.75	0.49
Lead Teacher Credential	CDA/ECE Course (n=6)	8.48	15.73	33.61	6.64	5.31	30.23	0.00
	AA (n=31)	7.85	20.08	28.07	4.18	5.46	33.58	0.77
	BA/BS (n=44)	6.88	20.61	23.77	4.84	6.60	36.48	0.82
	MA/ME (n=2)	4.73	18.93	34.02	16.86	0.00	21.89	3.55
	Missing Data (n=43)	7.62	22.35	26.69	4.85	6.61	31.21	0.67
PHLpreK Partner Agency	Phila SD (n=14)	8.09	19.63	31.10	2.75	4.21	32.85	1.38
	PHMC (n=50)	8.11	20.81	27.44	4.14	4.10	35.07	0.33
	UAC (n=62)	6.69	21.11	24.57	6.13	8.30	32.22	0.98
Council District	1 (n=5)	8.11	21.35	22.26	7.65	2.63	37.21	0.80
	2 (n=22)	8.40	20.51	25.28	5.11	2.67	37.64	0.39
	3 (n=24)	7.24	22.68	30.21	3.36	7.44	28.75	0.32
	4 (n=3)	6.20	13.18	29.46	9.88	1.74	39.53	0.00
	5 (n=9)	9.29	25.45	21.44	1.53	1.84	40.39	0.06
	6 (n=15)	8.81	18.22	28.78	6.10	5.07	32.83	0.19
	7 (n=15)	7.66	23.78	19.45	3.45	8.89	34.97	1.79
	8 (n=16)	6.31	17.45	31.13	6.91	8.69	27.13	2.38
	9 (n=9)	6.45	18.40	26.96	6.77	7.86	32.84	0.70
	10 (n=8)	3.50	22.18	24.54	4.28	10.41	34.81	0.29
Curriculum	Creative (n=77)	6.96	21.23	24.73	4.18	6.12	36.29	0.49
	Creative + another (n=11)	9.36	20.80	35.72	5.35	3.16	25.51	0.11
	Mother Goose Time (n=13)	8.14	18.79	26.56	8.61	9.72	27.67	0.51
	Other (n=23)	6.55	20.80	28.42	5.55	4.50	32.01	2.17

A second lens provided by the EduSnap is the percentage of time spent in different components of literacy. Figure 6 below presents the average for different components across the classrooms observed. While different programs may be stronger in different aspects of literacy, or preferably balanced across these, what is notable is all components of literacy are extremely low. Even very little time was spent reading to children on average across classes (on average, about 8 minutes per day).

Figure 6. Percentages of Time and Minutes Spent in Components of Literacy.

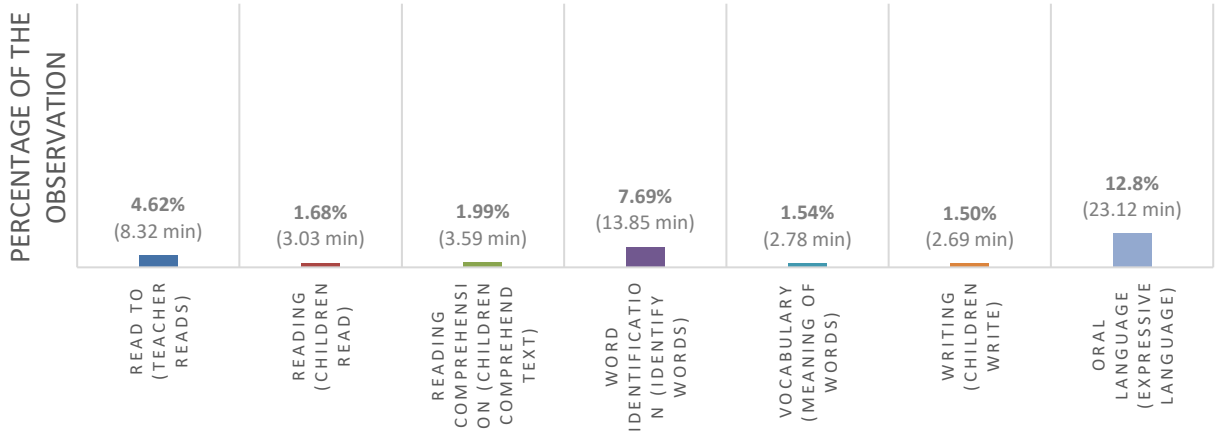


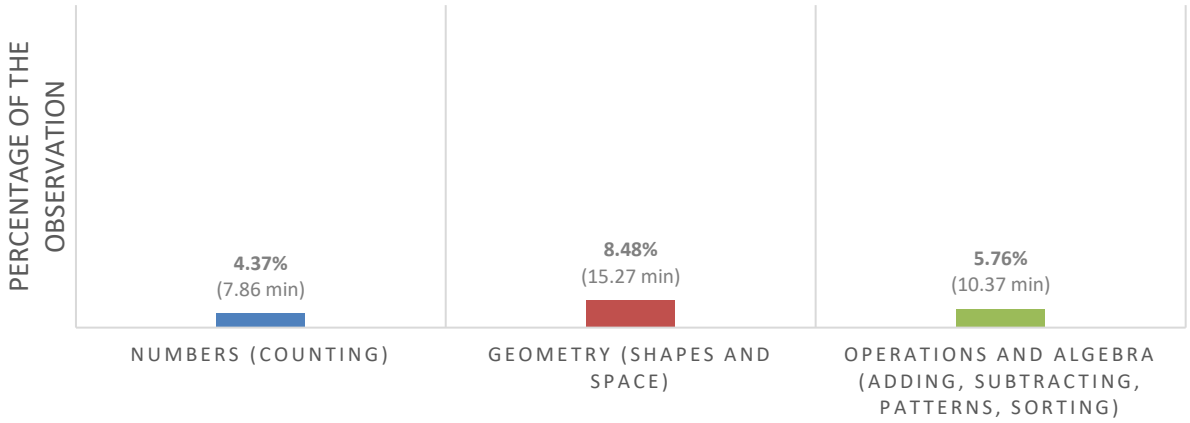
Table 5 shows this same percentages but for each subgroup of interest. Interestingly, one clear pattern here is that unrated classrooms seem to outperform other classrooms in several of the literacy components, followed by STAR 3 rated classrooms (except for reading to children). Oral language appears much higher in classrooms that have teachers with a CDA/ECE or MA/ME.

Table 5. Percentages of time spent in Components of Literacy by Subgroup, n=126.

		% of Time						
		Read-to	**Reading**	**Compreh.**	**Word ID**	**Vocabulary**	**Writing**	**Oral Language**
STAR Ratings	1 (n=16)	5.22	1.34	1.99	7.64	1.45	2.57	9.09
	2 (n=21)	5.65	1.10	2.11	6.49	1.04	1.13	13.68
	3 (n=58)	3.88	2.00	1.73	8.42	1.38	1.17	13.40
	4 (n=25)	4.77	1.65	2.19	6.87	2.40	1.84	12.76
	Unrated (n=6)	6.23	1.61	3.44	8.38	1.50	1.50	15.47
Number of PHL Contracted Slots	0-18 (n=25)	4.54	0.82	2.36	6.23	1.80	1.27	14.18
	19-28 (n=38)	5.14	1.79	2.58	8.20	1.18	1.21	13.93
	29-48 (n=30)	4.21	2.57	2.07	8.04	1.31	1.00	12.77
	49-90 (n=33)	4.47	1.39	0.97	7.89	1.98	2.45	10.66
Number of PHL preK classroom	1 (n=40)	5.26	1.64	2.80	8.04	1.55	0.70	15.09
	2 (n=40)	3.98	2.08	1.66	7.16	1.49	1.57	12.33
	3 (n=24)	6.03	1.42	1.84	5.29	2.03	1.91	10.59
	4 or more (n=22)	3.06	1.34	1.26	10.70	1.09	2.38	12.12
Lead Teacher Credential	CDA/ECE Course (n=6)	3.27	1.94	2.45	11.44	2.15	0.00	21.45
	AA (n=31)	5.05	1.70	1.70	7.93	1.00	1.43	11.77
	BA/BS (n=44)	4.83	2.01	2.46	7.84	1.89	2.11	11.75
	MA/ME (n=2)	3.25	0.00	0.59	9.17	1.78	0.00	37.28
	Missing Data (n=43)	4.35	1.38	1.74	6.80	1.49	1.19	12.46
PHLpreK Partner Agency	Phila SD (n=14)	5.04	1.83	1.92	5.59	0.96	0.79	11.80
	PHMC (n=50)	4.41	1.54	2.22	8.04	2.13	1.62	13.62
	UAC (n=62)	4.70	1.76	1.82	7.89	1.20	1.55	12.45
Council District	1 (n=5)	5.25	4.57	2.05	9.13	0.34	1.71	8.33
	2 (n=22)	5.22	1.49	2.39	6.29	1.21	0.90	18.46
	3 (n=24)	4.77	1.63	2.57	9.98	2.15	2.37	14.08
	4 (n=3)	7.95	2.33	4.07	8.33	1.36	0.00	14.92
	5 (n=9)	2.99	1.46	1.72	7.70	1.08	1.65	10.56
	6 (n=15)	4.08	2.10	1.49	6.75	1.56	1.07	14.14
	7 (n=15)	6.59	0.95	1.98	6.43	0.91	0.52	5.52
	8 (n=16)	2.53	1.62	0.68	7.78	1.59	2.76	16.32
	9 (n=9)	4.54	1.41	1.85	7.67	2.11	1.21	8.37
	10 (n=8)	4.42	1.50	2.43	7.42	2.35	1.21	8.84
Curriculum	Creative (n=77)	4.10	1.93	1.83	7.61	1.53	0.85	12.89
	Creative + another (n=11)	5.99	1.23	3.37	12.09	2.78	1.39	19.95
	Mother Goose Time (n=13)	2.13	1.20	1.39	8.56	0.60	2.64	13.60
	Other (n=23)	7.02	1.47	2.12	5.45	1.52	3.25	9.40

Similarly, the EduSnap looks at the percentage of time spent in three specific components of math: numbers, geometry and operations and algebra (Figure 7). Percentages for these resembled those of literacy, with only a slightly higher percentage observed for time spent in geometry.

Figure 7. Percentages of Time and Minutes Spent in Components of Math.



Percentages of time spent in components of math for each subgroup of interest are shown in Table 6. Higher rated classrooms spend slightly more time on numbers and geometry, but not necessarily on algebra, compared to lower rated classrooms. The percentage of time spent on algebra varied from 1.6% to 12.1% across different types of settings, much more than the variation observed in number or geometry.

Table 6. Percentages of time spent in components of math by subgroups, n=126.

		% of Time		
		Numbers	Geometry	Algebra
STAR Ratings	1 (n=16)	3.11	5.03	7.03
	2 (n=21)	3.92	7.18	5.74
	3 (n=58)	4.06	9.42	6.06
	4 (n=25)	5.99	9.52	5.19
	Unrated (n=6)	5.48	8.81	1.61
Number of PHL Contracted Slots	0-18 (n=25)	4.04	7.50	7.21
	19-28 (n=38)	4.10	7.58	6.94
	29-48 (n=30)	3.63	9.51	5.80
	49-90 (n=33)	5.60	9.31	3.28
Number of PHL preK classrooms	1 (n=40)	4.28	6.73	6.83
	2 (n=40)	3.89	9.53	6.58
	3 (n=24)	4.83	9.68	5.32
	4 or more (n=22)	4.89	8.49	2.76
Lead Teacher Credential	CDA/ECE Course (n=6)	5.82	4.19	6.54
	AA (n=31)	4.11	7.35	6.07
	BA/BS (n=44)	4.61	8.61	5.64
	MA/ME (n=2)	6.21	9.17	3.85
	Missing Data (n=43)	4.02	9.72	5.65
PHLpreK Partner Agency	Phila SD (n=14)	3.21	7.34	10.59
	PHMC (n=50)	5.07	8.60	5.95
	UAC (n=62)	4.05	8.65	4.49
Council District	1 (n=5)	3.77	8.45	12.56
	2 (n=22)	3.03	9.78	3.31
	3 (n=24)	4.40	6.99	6.35
	4 (n=3)	6.78	2.71	1.55
	5 (n=9)	3.82	7.19	5.60
	6 (n=15)	4.99	8.27	12.12
	7 (n=15)	3.65	8.93	5.44
	8 (n=16)	5.89	10.54	2.61
	9 (n=9)	3.13	7.92	1.73
	10 (n=8)	6.35	9.42	6.92
Curriculum	Creative (n=77)	4.13	9.37	4.35
	Creative + another (n=11)	7.06	5.78	5.51
	Mother Goose Time (n=13)	4.77	4.91	8.24
	Other (n=23)	3.85	9.25	8.83

The EduSnap also takes into account the importance of Content Area Balance, and therefore captures the percentage of time spent in the different content areas (Figure 8). Even though the literacy activities were low, they are higher than other areas. Science and gross motor are the two content areas with the lowest percentages of time observed.

Figure 8. Percentages of Time and Minutes Spent in Content Areas.

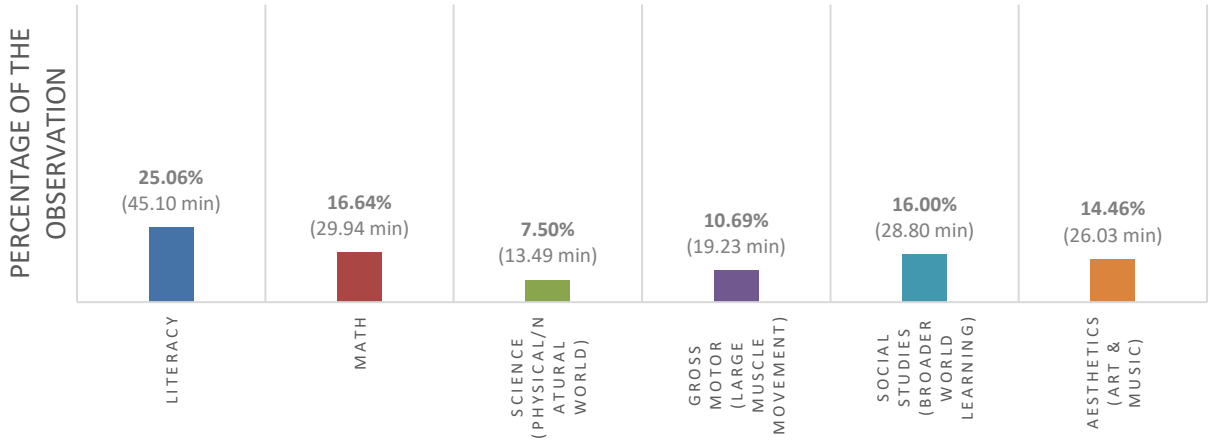


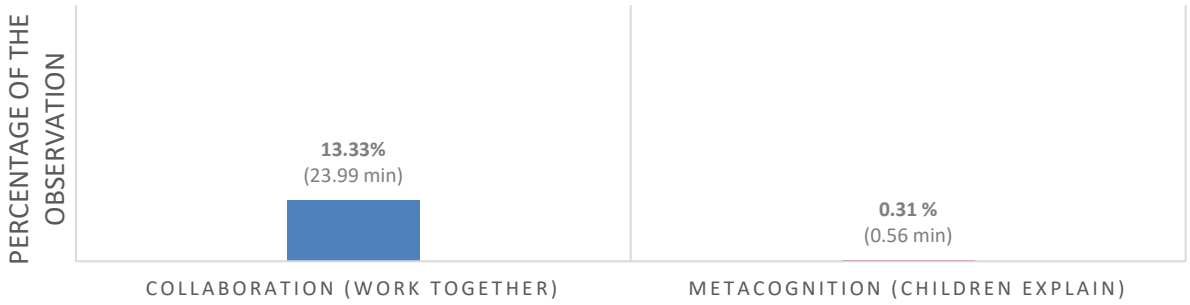
Table 7 shows the percentage of time spent in the different content areas for the various subgroups of interest. STAR 4 classrooms spend the highest percentages of time in math and science content, followed by STAR 3 classrooms. On average, this difference in content amounts to about 14 or 15 minutes per day for STAR 4 classrooms compared to STAR 1 or STAR 2 classrooms. Additionally, Phila SD programs engaged in the greatest amount of math and science content, followed by PHMC programs, compared to the other partner agencies. These content areas were also more frequent in classrooms with teachers with a CDA/ECE or MA/ES.

Table 7. Percentages of time spent in various content areas by subgroups, n=126.

		% of Time					
		Literacy	Math	Science	Gross Motor	Social Studies	Aesthetics
STAR Ratings	1 (n=16)	23.76	14.05	6.34	10.94	16.84	14.78
	2 (n=21)	23.86	14.97	5.71	10.13	15.78	13.56
	3 (n=58)	25.84	17.43	7.03	10.61	16.45	14.46
	4 (n=25)	24.64	18.16	10.57	10.48	15.88	15.74
	Unrated (n=6)	26.96	15.04	8.38	13.75	9.99	10.96
Number of PHL Contracted Slots	0-18 (n=25)	24.98	16.97	10.96	11.39	15.02	14.42
	19-28 (n=38)	26.37	16.51	6.79	12.46	15.28	15.70
	29-48 (n=30)	24.64	16.60	5.93	8.58	20.08	14.82
	49-90 (n=33)	23.98	16.56	7.18	10.06	13.76	12.71
Number of PHL preK classrooms	1 (n=40)	27.18	15.53	8.83	11.87	16.04	14.63
	2 (n=40)	24.35	17.75	8.33	9.76	17.67	15.21
	3 (n=24)	22.10	17.99	8.11	9.73	14.33	14.24
	4 or more (n=22)	25.66	15.15	2.78	11.25	14.69	13.02
Lead Teacher Credential	CDA/ECE Course (n=6)	32.58	15.42	10.11	12.05	12.87	20.74
	AA (n=31)	25.35	16.05	7.76	11.45	16.31	15.61
	BA/BS (n=44)	24.99	16.75	5.72	9.92	16.45	15.28
	MA/ME (n=2)	45.86	15.38	2.96	20.41	16.27	13.02
	Missing Data (n=43)	22.94	17.16	8.97	10.27	15.71	12.02
PHLpreK Partner Agency	Phila SD (n=14)	23.55	17.88	10.92	7.88	19.26	12.92
	PHMC (n=50)	25.91	17.55	8.57	11.12	16.09	14.93
	UAC (n=62)	24.70	15.60	5.82	10.98	15.17	14.43
Council District	1 (n=5)	25.00	20.21	5.82	10.05	11.87	16.67
	2 (n=22)	27.08	15.25	5.03	12.16	15.96	14.69
	3 (n=24)	28.51	15.81	5.76	10.72	16.77	15.74
	4 (n=3)	28.29	9.69	4.84	14.53	19.57	13.18
	5 (n=9)	21.95	14.12	8.97	15.27	18.89	9.73
	6 (n=15)	26.76	22.11	10.33	8.08	18.60	15.97
	7 (n=15)	18.14	16.55	5.08	6.63	15.52	11.95
	8 (n=16)	27.99	17.79	10.28	11.52	13.07	14.66
	9 (n=9)	22.11	11.88	11.37	12.14	13.80	17.06
	10 (n=8)	19.26	18.69	8.63	9.63	15.83	12.98
Curriculum	Creative (n=77)	24.10	15.77	5.99	11.47	16.39	13.85
	Creative + another (n=11)	35.19	15.94	10.32	17.38	13.74	14.12
	Mother Goose Time (n=13)	25.50	16.47	7.96	8.10	16.80	17.95
	Other (n=23)	23.73	20.06	10.95	6.73	15.72	15.45

In terms of student learning approaches (Figure 9), the observation showed that only 13% of the time was used in collaborative approaches and no time in metacognition.

Figure 9. Percentages of Time and Minutes Spent in Student Learning Approaches.



Student learning approaches by type of setting are shown in Table 8. Most importantly, the absence of metacognition is true for all types of providers, star levels, types of teachers, agencies and curriculum approaches. Collaborative approaches vary between 10% and 21%, and are higher classrooms with teachers with a CDA/ECE or MA/ME.

Table 8. Percentages of time spent in student learning approaches by subgroups, n=126.

		% of Time	
		Collaboration	Metacognition
STAR Ratings	1 (n=16)	15.68	0.18
	2 (n=21)	11.45	0.20
	3 (n=58)	13.80	0.30
	4 (n=25)	11.99	0.49
	Unrated (n=6)	14.50	0.32
Number of PHL Contracted Slots	0-18 (n=25)	15.65	0.43
	19-28 (n=38)	14.33	0.25
	29-48 (n=30)	11.61	0.31
	49-90 (n=33)	12.03	0.29
Number of PHL preK classrooms	1 (n=40)	14.38	0.38
	2 (n=40)	13.99	0.28
	3 (n=24)	11.00	0.34
	4 or more (n=22)	12.75	0.19
Lead Teacher Credential	CDA/ECE Course (n=6)	21.35	0.41
	AA (n=31)	14.47	0.21
	BA/BS (n=44)	13.78	0.43
	MA/ME (n=2)	18.93	0.00
	Missing Data (n=43)	10.72	0.26
PHLpreK Partner Agency	Phila SD (n=14)	13.13	0.33
	PHMC (n=50)	12.46	0.40
	UAC (n=62)	14.09	0.23
Council District	1 (n=5)	14.38	0.57
	2 (n=22)	17.25	0.25
	3 (n=24)	9.78	0.25
	4 (n=3)	16.28	0.00
	5 (n=9)	14.57	0.13
	6 (n=15)	12.28	0.53
	7 (n=15)	10.28	0.24
	8 (n=16)	17.04	0.34
	9 (n=9)	11.44	0.26
	10 (n=8)	13.05	0.50
Curriculum	Creative (n=77)	14.03	0.29
	Creative + another (n=11)	13.42	0.59
	Mother Goose Time (n=13)	12.36	0.28
	Other (n=23)	12.23	0.29

Figure 10 presents the percentages of time in didactic and scaffolded teaching approaches. Classrooms in the sample tend to weigh toward didactic approaches over scaffolded approaches.

Figure 10. Percentages of Time and Minutes Spent in Teaching Approaches.

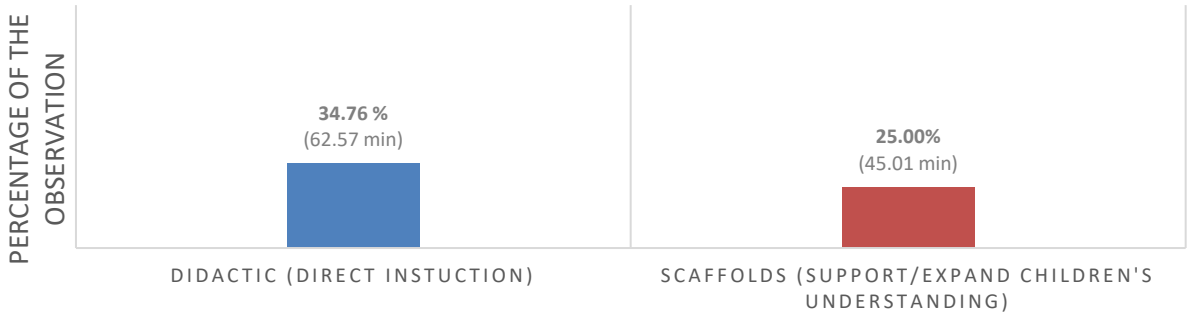


Table 9 shows the percentage of time spent in didactic or scaffolded approaches by subgroups of interest. Higher percentage of scaffolding or didactic approaches are not observed in higher star level classrooms. Higher scaffolding is, however, present in classrooms with teachers with a CDA/ECE or an MA/ME and this difference is equivalent to 18 to 36 minutes.

Table 9. Percentages of time spent in Teaching Approaches by Subgroup, n=126.

		% of Time	
		Didactic	Scaffolds
STAR Ratings	1 (n=16)	36.11	20.35
	2 (n=21)	37.36	22.24
	3 (n=58)	34.48	27.41
	4 (n=25)	32.74	22.17
	Unrated (n=6)	33.40	36.63
Number of PHL Contracted Slots	0-18 (n=25)	34.86	29.23
	19-28 (n=38)	35.25	25.28
	29-48 (n=30)	33.41	24.21
	49-90 (n=33)	35.38	22.25
Number of PHL preK classrooms	1 (n=40)	33.61	29.55
	2 (n=40)	36.32	24.15
	3 (n=24)	34.94	18.55
	4 or more (n=22)	33.85	25.28
Lead Teacher Credential	CDA/ECE Course (n=6)	33.40	44.74
	AA (n=31)	32.83	25.60
	BA/BS (n=44)	34.56	22.49
	MA/ME (n=2)	25.44	62.13
	Missing Data (N=43)	36.98	22.76
PHLpreK Partner Agency	Phila SD (n=14)	39.06	23.22
	PHMC (n=50)	32.29	28.91
	UAC (n=62)	35.81	22.20
Council District	1 (n=5)	43.49	21.92
	2 (n=22)	30.70	25.45
	3 (n=24)	38.12	30.34
	4 (n=3)	41.67	28.68
	5 (n=9)	31.55	24.68
	6 (n=15)	32.67	30.08
	7 (n=15)	37.79	12.11
	8 (n=16)	35.13	30.53
	9 (n=9)	31.95	20.19
	10 (n=8)	31.88	18.05
Curriculum	Creative (n=77)	32.37	23.95
	Creative + another (n=11)	30.86	43.05
	Mother Goose Time (n=13)	41.05	26.56
	Other (n=23)	40.31	20.01

An important piece in curriculum management, is the integration of curriculum across areas (Figure 11). In pre-K classrooms, this means that teachers are successful in carrying out multiple content areas simultaneously. In the sample, classrooms evidenced no integration 34% of the time (i.e. children were not receiving any content), while there was at least one content area occurring 43% of the time. Higher levels of integration occurred about a quarter of the time.

Figure 11. Percentages of Time and Minutes Spent in Curriculum Integration

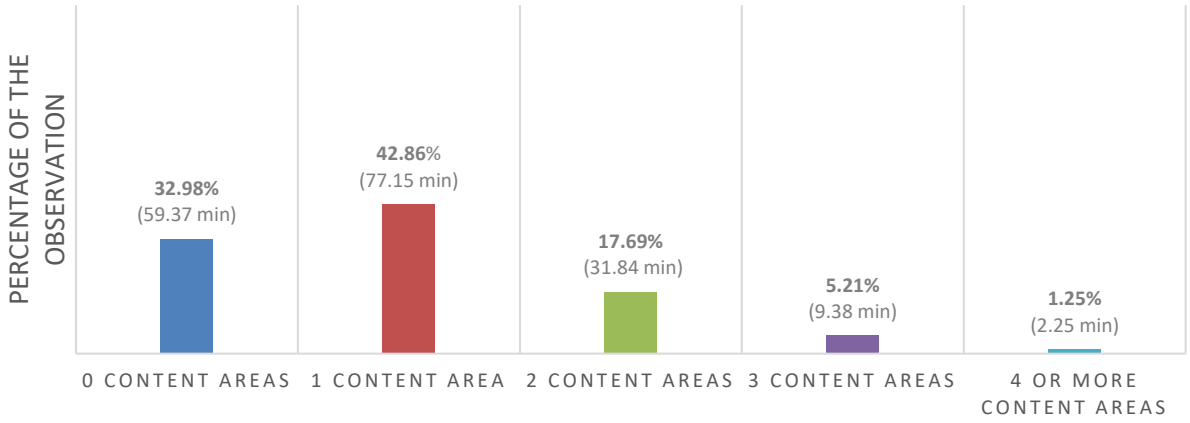


Table 10 describes the percentages of curriculum integration across our subgroups of interest. The distribution resembles that for the overall group. A few notable trends are that higher STAR rating classrooms had slightly higher percentages of integration of two or three areas. Further, more integration was found in classrooms with teachers with a CDA/ECE or an MA/ME (though again, it is important to note that teacher qualifications were missing for about 40% of our sample). This difference was about 18 minutes of integrated curriculum (in one or more areas), on average.

Table 10. Percentages of time spent in Curriculum Integration by Subgroup, n=126.

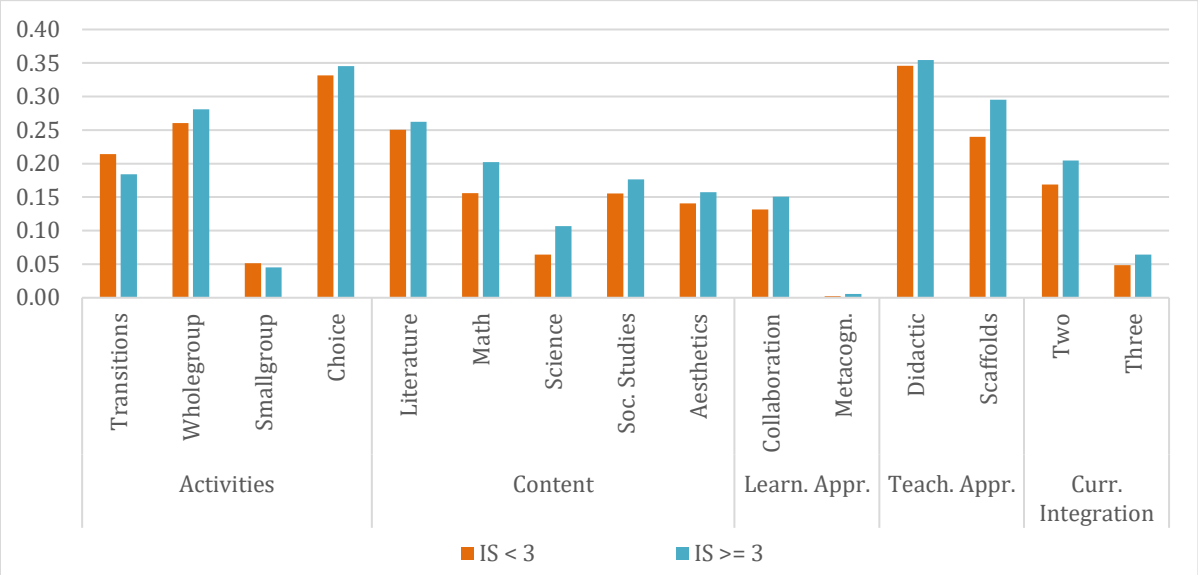
		% of Time				
		None	1	2	3	4 or more
STAR Ratings	1 (n=16)	34.34	44.66	15.94	3.69	1.38
	2 (n=21)	36.84	41.49	15.12	5.08	1.47
	3 (n=58)	32.86	41.91	18.96	5.12	1.15
	4 (n=25)	29.83	43.53	19.00	6.38	1.26
	Unrated (n=6)	30.50	49.73	13.00	5.69	1.07
Number of PHL Contracted Slots	0-18 (n=25)	32.28	42.38	18.37	5.72	1.25
	19-28 (n=38)	31.67	42.99	17.84	6.04	1.46
	29-48 (n=30)	32.70	42.45	18.55	4.73	1.57
	49-90 (n=33)	35.31	43.46	16.20	4.30	0.72
Number of PHL preK classrooms	1 (n=40)	31.45	42.55	18.22	6.22	1.56
	2 (n=40)	32.08	42.63	18.87	5.06	1.35
	3 (n=24)	33.64	43.94	16.79	4.78	0.86
	4 or more (n=22)	36.8	42.67	15.53	4.07	0.93
Lead Teacher Credential	CDA/ECE Course (n=6)	24.21	46.17	21.65	6.54	1.43
	AA (n=31)	30.95	45.9	17.31	4.88	0.96
	BA/BS (n=44)	33.95	42.16	16.63	5.65	1.61
	MA/ME (n=2)	24.26	38.76	26.04	10.65	0.3
	Missing Data (N=43)	35.07	41.12	18.12	4.57	1.12
PHLpreK Partner Agency	Phila SD (n=14)	31.72	43.89	17.72	5.96	0.71
	PHMC (n=50)	31.58	41.9	19.43	5.56	1.53
	UAC (n=62)	34.44	43.42	16.25	4.74	1.15
Council District	1 (n=5)	32.53	43.84	16.32	6.05	1.26
	2 (n=22)	30.00	47.64	16.07	5.28	1.01
	3 (n=24)	34.17	38.17	19.79	5.56	2.32
	4 (n=3)	27.71	50.58	14.73	5.81	1.16
	5 (n=9)	33.46	42.88	18.70	3.88	1.08
	6 (n=15)	30.58	39.23	22.00	6.60	1.60
	7 (n=15)	42.36	40.25	12.70	4.29	0.40
	8 (n=16)	29.54	44.92	20.14	4.95	0.45
	9 (n=9)	32.84	46.13	15.85	4.28	0.89
	10 (n=8)	33.17	44.79	14.91	5.35	1.78
Curriculum	Creative (n=77)	33.75	43.66	16.41	5.12	1.07
	Creative + another (n=11)	26.04	40.32	23.74	7.33	2.57
	Mother Goose Time (n=13)	35.91	37.85	19.62	5.04	1.57
	Other (n=23)	30.54	45.00	18.70	4.64	1.13

What do high-quality classrooms have in common?

Lastly, for this report we looked closely at a subset of classrooms that scored highest on the *Instructional Support* domain of the CLASS given the relationship with this domain and children's learning. We attempted to identify commonalities that existed across these classrooms, to understand why they may have exemplified higher quality classroom practices, and also understand the balance in activities, content, integration in curriculum, and teaching and learning in higher quality classrooms in contrast to lower quality classrooms.

Figure 12 compares average EduSnap percentages for classrooms with instructional supports below and above the threshold of 3. There are three main things that come out of looking at the data this way: (a) higher quality classrooms are those that spend on average slightly less time on transitions (about 5 less minutes); (b) they also provide more content across all areas (which adds up to a difference of almost 25 minutes), particularly in math and science; (c) they use more scaffolded teaching approaches (average difference is equivalent to 10 minutes); and (d) they integrate more (difference in integration across 2 or 3 content areas is equivalent to 9 minutes). While not illustrated here, we also found differences in content and teaching approaches for classrooms that scored above a 2.3 in IS, compared to those that scored below a 2.3 (which is the median score for the classrooms in the sample). This suggests a positive association between usage of time in content and teaching approaches and CLASS IS at various levels of quality.[1]

Figure 12. EduSnap average percentages grouped by CLASS Instructional support levels



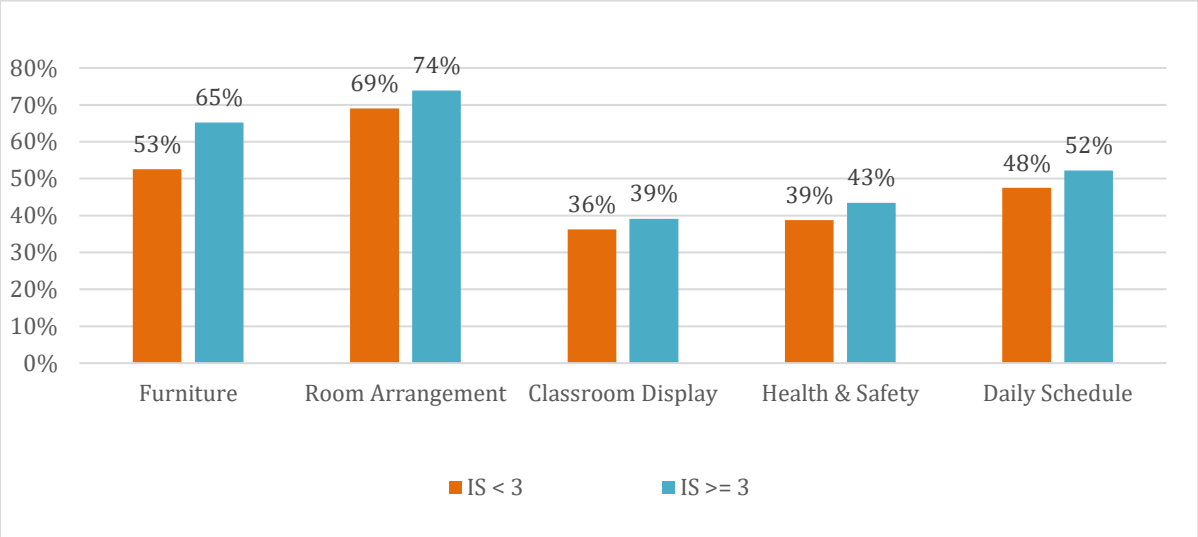
We also examined whether classrooms that scored above a 3.0 on CLASS IS were more likely to be implementing a specific curriculum, and we found that curriculum did not make a difference in IS scores. Further, simple regression analyses were used to examine associations between CLASS IS and teacher qualification levels, curriculum, and star levels. We found minimal differences except for positive associations between master's degree and CLASS IS, however this should be interpreted with caution because there were only two teachers in the sample with a Master's degree.

In addition, we looked at whether classrooms with higher CLASS IS scores differed from lower CLASS IS classrooms in terms of the indicators included in the checklist. The environment may be in facilitating interactions that make this support easier or harder. The checklist focused on indicators around furniture, room arrangement, classroom display, health

[1] An estimation of the association between CLASS IS scores and the probability of being scored above 3 in CLASS IS and the different uses of time in the classrooms confirmed a positive association between higher content, particularly in math and science, and CLASS IS. This estimation also showed a positive association between metacognitive activities as minimal as the were, and CLASS IS scores for classrooms in PHL.

and safety, and daily schedules. Figure 13 below show the percentage of low CLASS IS and high CLASS IS that met all indicators in the checklist under each of these areas, and the percentage of classroom meeting all indicators overall. The group of classrooms that are scoring higher on instructional support are in fact on average meeting all indicators under some of the areas on the checklist at a slightly higher rate.

Figure 13. Checklist percentages met by area by CLASS Instructional support levels



Discussion of Findings

This is the first report of classroom quality for the evaluation of Philadelphia's PreK program. As such, and given that PHLpreK has been creating partnership with a great number of existing programs, it should be seen as a baseline on quality, as well as a compass or GPS to design professional development and feed into any continuous improvement strategies early on. Classroom observations in pre-K will continue going forward for the length of the research study. Pre-K classrooms in these programs are averaging moderate levels of quality as measured by the CLASS Emotional Support and Classroom Organization, but very low levels of quality on the CLASS Instructional support domain. That is, on average classrooms are adequately nurturing and safe environments for children and adequately structured and organized. However, teachers' use of strategies and techniques for scaffolding and expanding children's learning and language are infrequent.

The EduSnap showed that classrooms are generally effective at implementing a variety of activity settings. A third of the day is spent in choice time with many learning opportunities; however, in contrast another 20% of the day is spent in transitions with offer few opportunities for learning for young children. Teachers are effective at carrying out multiple content areas for about a fourth of the day, but no content occurs for another third of the day. Further, while children experience a balance of didactic and scaffolded interactions, children are rarely asked to explain or justify their thinking through metacognitive processes.

Children's experiences across subgroups were similar with a few exceptions. First, children in STAR level 1 programs spend more time in whole groups and less time in choice

compared to children in other programs. In addition, children in programs with just one PHLpreK classroom experience more scaffolded interactions with teachers and more collaboration with one another. Lastly, children in STAR level 1 and 2 programs spend more time in didactic interactions compared to children in STAR level 3 and 4 programs. In contrast, children in STAR 3 and 4 programs spend more time in math and literacy content and in multiple content areas simultaneously, compared to children in STAR 1 and 2 programs.

We also took a closer look at the subset of classrooms that scored above a 3.0 on CLASS instructional support and found a few notable characteristics. Specifically, children in classrooms that score above 3.0 spend more time in every activity setting, except for transitions. They also spend more time in every content area and teaching approach. In particular, this association is stronger for math and science and integration of content. Overall, this shows that classrooms are more consistent and richer in all of the content and interactions that children are exposed to when they score a 3.0 or higher on CLASS IS. The data indicates that strengthening content, integration, and metacognition in the classroom would also go hand in hand with strengthening instructional support. There are also positive associations with CLASS IS scores and teacher with a Master's degree. However, it is difficult to draw conclusions around this finding given that very few teachers in our sample have a Master's degree. We do not find any positive associations between high-scoring CLASS IS classrooms and curriculum or star level, but we do observe stronger infrastructure does support higher CLASS IS levels.

Acknowledgments

We are grateful to the William Penn Foundation who funded the work of NIEER that made this research possible (grant No. 16-17). We are also grateful for the partnership with the City of Philadelphia and PHMC, which were key in supporting our work with program providers throughout the City. Finally, we are thankful for all the programs and teachers that opened their doors to us. The authors of this report are also grateful to Hebbah El-Moslimany for her coordination work with the centers and Ruiya Zhang for her statistical analyses.

References

Barnett, W. S. (2008). Preschool education and its lasting effects: Research and policy implications. Boulder and Tempe: Education and the Public Interest Center & Education Policy Research Unit. Retrieved September 20, 2017, from http://nepc.colorado.edu/files/PB-Barnett-EARLY-ED_FINAL.pdf

Barnett, W. S. (2011). Effectiveness of early educational intervention. Science, 333, 975-978.

Barnett, W. S., & Nores, M. (2015). Investment and productivity arguments for ECCE. Investing against Evidence, 73.

Burchinal, M., Vandergrift, N., Pianta, R., & Mashburn, A., (2009). Threshold analysis of association between child care quality and child outcomes for low-income children in pre-kindergarten programs. Early Childhood Research Quarterly. 25, 166-176.

Burchinal, M., Kainz, K., & Cai, Y. (2011). How well do our measures of quality predict child outcomes? A meta-analysis and coordinated analysis of data from large-scale studies of early childhood settings. In M. Zaslow, I. Martinez-Beck, K. Tout, & T. Halle (Eds.), Quality measurement in early childhood settings (pp. 11–31). Baltimore, MD: Brookes.

Burchinal, M., Vernon-Feagans, L., Vitiello, V., & Greenberg, M. (2014). Thresholds in the Association between child care quality and child outcomes in rural preschool children. Early Childhood Research Quarterly, 29, 41–51. doi:10.1016/j.ecresq.2013.09.004

Camilli, G., Vargas, S., Ryan, S., & Barnett, W.S. (2010). Meta-analysis of the effects of early education interventions on cognitive and social development. Teachers College Record, 112, 579-620.

Ceci, S. J., & Papierno, P. B. (2005). The Rhetoric and Reality of Gap Closing: When the "Have-Nots" Gain but the "Haves" Gain Even More. American Psychologist, 60(2), 149-160.

Duncan, G, and Murnane, R. (2011). Introduction: The american dream, then and now. In eds. Greg J. Duncan, Richard J. Murnane, (Eds.), *Whither Opportunity: Rising Inequality, Schools, and Children's Life Chances* (pp. 3–26). New York: Russell Sage Foundation and Spencer Foundation.

Friedman-Krauss, A., Barnett, S., & Nores, M. (2016). How much can high-quality universal pre-K reduce achievement gaps? Washington, DC: Center for American Progress and National Institute for Early Education Research. Retrieved from https://cdn.americanprogress.org/wp-content/uploads/2016/04/01115656/NIEER-AchievementGaps-report.pdf

Hamre, B., Hatfield B. E., Pianta R. C., & Jamil F. (2014). “Evidence for General and Domain-Specific Elements of Teacher-Child Interactions: Associations with Preschool Children’s Development,” Child Development, 85, 1257–1274.

Hatfield, B. E., Lower, J. K., Cassidy, D., Faldowski, R. (2015). Inequities in access to quality early care and education: Associations with funding and community context. Early Childhood Research Quarterly Special Issue: Quality Rating and Improvement Systems (QRIS) as Change Agents, 30, 316-326. doi:10.1016/j.ecresq.2014.01.00hjm

Hatfield, B. E., Burchinal, M. R., Pianta, R. C., & Sideris, J. (2016). Thresholds in the association between quality of teacher–child interactions and preschool children's school readiness skills. Early Childhood Research Quarterly, 36, 561–571. doi:10.1016/j.ecresq.2015.09.005

Hestenes, L. L., Kintner-Duffy, V., Wang, Y. C., La Paro, K., Mims, S. U., Crosby, D., & Cassidy, D. J. (2015). Comparisons among quality measures in child care settings: Understanding the use of multiple measures in North Carolina's QRIS and their links to social-emotional development in preschool children. Early Childhood Research Quarterly, 30, 199-214.

Martinez-Beck, I. (2011). Introduction: why strengthening the measurement of quality in early childhood settings has taken on new importance, in: M. Zaslow, I. Martinez-Beck, K. Tout & T. Hall (Eds). *Quality Measurement in Early Childhood Settings*. Baltimore, MD: Brookes.

N. Aikens, L. Tarullo, L. Hulsey, C. Ross, J. West, Y. Xue (2010). ACF-OPRE Report: A Year in Head Start: Children, Families, and Programs. Washington, DC. U.S. Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research and Evaluation.

Office of Head Start, U.S. Department of Health and Human Services, Administration for Children and Families. (2015). A National Overview of Grantee CLASS Scores in 2015. Washington, DC.

Philadelphia Commission on Universal Pre-Kindergarten. (2016). *Final Recommendations Report*. Retrieved September 20, 2017, from http://www.phila.gov/universalprek/Documents/Recommendations%20Report.pdf

Pianta, R. C., & Hamre, B. K. (2009). Conceptualization, measurement, and improvement of classroom processes: Standardized observation can leverage capacity. Educational Researcher, 38 (2009), pp. 109-119, 10.3102/0013189X09332374

Weiland, C., & Yoshikawa, H. (2013). Impacts of a prekindergarten program on children's mathematics, language, literacy, executive function, and emotional skills. Child Development, 84(6), 2112-2130.

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.

Yoshikawa, H., Wieland, C., Brooks-Gunn, J., Burchinal, M. R., Espinosa, L. M., & Gormley, W. T. (2013). Investing in our future: The evidence base on preschool education. Society for Research in Child Development.

Table I. Classroom Observation Checklist

	Indicator	Yes/No	Notes
Furniture			
1	Routine care furniture is convenient to use.	□Yes □No	
2	Woodwork bench, sand/water table, or easel is used.	□Yes □No	
3	Some soft furnishings (i.e. carpet area, soft chair) accessible to children (at least 2)	□Yes □No	
4	Most soft furnishings are clean and in good repair.	□Yes □No	
Room Arrangement			
5	Quiet and active centers placed to not interfere with one another	□Yes □No	
6	Space is arranged so most activities are not interrupted.	□Yes □No	
7	At least five different interest centers provide a variety of learning experiences.	□Yes □No	
8	Centers are organized for independent use by children (open, labeled shelves).	□Yes □No	
Classroom Display			
9	Much of the display relates closely to current activities and children in group.	□Yes □No	
10	Much of the display is work done by the children.	□Yes □No	
11	Individualized children's work predominates.	□Yes □No	
Health and Safety			
12	The classroom has running water.	□Yes □No	
13	Child restrooms are located in the classroom.	□Yes □No	
14	No major safety hazards indoors	□Yes □No	
15	Adequate supervision to protect children's safety indoors	□Yes □No	
Daily Schedule			
16	Written schedule is posted in room.	□Yes □No	
17	At least one indoor and one outdoor play period occurs daily.	□Yes □No	
18	Schedule provides balance of structure and flexibility.	□Yes □No	
19	A variety of play activities are scheduled each day.	□Yes □No	
20	This classroom is similar in materials and climate to other observed classrooms in this center.	□Yes □No □N/A	

FURNITURE & MATERIALS	**Check if present**
Cubbies, for each child	
Circle Time Carpet	
Child-sized Tables	
Child-sized Chairs	
Shelving Unit for Blocks	

Blocks for children's use	
Shelving Unit for Book Area	
Books for children's use	
Shelving Unit for Art	
Art materials for children's use	
Shelving Units for Dramatic Play Area	
Dramatic play materials for children's use	
Shelving Unit for Manipulatives	
Manipulatives for children's use	
Shelving unit for Nature/science	
Nature/science material's for children's use	
Shelving unit for music area	
Instruments for children's use	
Computer Stations	
Chairs for Computer Station	
Bean bag chair	
Easel	
Drying Rack	
Sand/Water Table	
Workbench	
Play Table and Chair Set	
Child-sized Sink	
Child-sized Stove	
Child-sized Refrigerator	
Child-sized Drawers	