

BPS K1DS: Piloting the Boston Public Schools'
Prekindergarten Model in Community-Based
Organizations



Final Report
February 2016

This page was intentionally left blank.

Preface

This report summarizes the process and results of the K1 in Diverse Settings (K1DS) pilot. The K1DS pilot project was a 2.5-year collaboration between the Boston Public Schools (BPS) and ten community-based organizations which offered early care and education services to 4-year-olds in high-poverty areas of Boston, MA. The pilot study introduced the hallmarks of the BPS prekindergarten (K1) model—developmentally appropriate, research-based language, literacy, and mathematics curricula supported by regular one-on-one coaching and training—to the participating classrooms. The findings summarized in this report provide an important entry point into understanding the context, nature, and quality of practices in community-based prekindergarten programs. This report also describes changes in classroom instructional quality and a range of other adult and child outcomes that were investigated during the study period. Because this pilot project was not meant to evaluate the causal impact of the BPS K1 model in CBO classrooms and, as a result, did not include an experimental design, these findings should be viewed as descriptive and should be used only for hypothesis generating purposes rather than evidence that any one component of the K1DS program *caused* any observed change in classroom quality or child performance on assessments.

The pilot program was designed by Dr. Jason Sachs in consultation with his team at the BPS Department of Early Childhood Education and community partners. The evaluation study was co-designed by Dr. Christina Weiland at the University of Michigan and Dr. Monica Yudron at the University of Massachusetts Boston. The authors would like to thank the directors, teachers, parents, and children at each of the participating programs as well as the BPS-based K1DS team for their generous participation in all evaluation activities. The authors would also like to thank Abt Associates and the Wellesley Centers for Women for collecting spring 2015 quality data. Finally, the authors dedicate this report to the memory of Michelle High-McKinnon whose profound commitment to the well-being of Boston's most vulnerable children manifested in many ways—including forming and maintaining the strong, collaborative relationships which formed the backbone of the BPS K1DS pilot.

Glossary

- Building Blocks- Building Blocks prekindergarten mathematics curriculum
- BPS- Boston Public Schools
- CBO-Community-based Organizations
- CLASS-Classroom Assessment Scoring System (Pianta, La Paro, & Hamre, 2008)
- COEMET- Classroom Observation of Early Mathematics Environment and Teaching (Sarama & Clements, 2009)
- ELLCO- Early Language and Literacy Classroom Observation (Smith, Dickinson, Sangeorge, & Anastasopoulos, 2002)
- K1-Prekindergarten classrooms in the Boston Public Schools
- CBO K1DS-K1 in Diverse Settings. Refers to the adaptation and implementation of the BPS K1 model in community-based prekindergarten classrooms
- OWL-Opening the World of Learning prekindergarten language, literacy, and socio-emotional curriculum
- SMA- refers to a specific mathematics activity, an in-depth teacher-led mathematics lesson typically delivered in small groups

Executive Summary

Background

School readiness gaps between advantaged and disadvantaged children are large at kindergarten entry and widen as children progress through elementary school (Duncan & Magnuson, 2011). Repeatedly, researchers have found that high-quality preschool programs can be an effective tool for reducing these gaps. But, on average in the United States, instructional quality—the type of quality that appears most consequential for children’s school readiness—is low (Yoshikawa et al., 2013). In the 2009-2010 school year, the Boston Public Schools (BPS) prekindergarten model (called K1) had the highest level of instructional quality of any evaluated large-scale prekindergarten in the U.S. (Weiland, Ulvestad, Sachs, & Yoshikawa, 2013). BPS K1 has proven positive impacts on children’s language, literacy, mathematics, and executive function skills at kindergarten entry (Weiland & Yoshikawa, 2013). The K1 model combines two evidence-based curricula in prekindergarten classrooms—one focused on language and literacy instruction and the other on mathematics instruction—with regular coaching and training supports for teachers.

Boston K1DS (K1 in Diverse Settings) was a 2.5-year pilot demonstration project to expand the nationally recognized BPS K1 model to 14 community-based preschool classrooms. This partnership between BPS, Thrive in 5, and community-based organizations (CBO) aimed to build a high quality, private- and city-funded network of early childhood pre-K classrooms to expand access for families, close the achievement gap, and improve academic outcomes for Boston children living in the Circle of Promise and East Boston. The pilot began in January 2013 and concluded in June 2015. This report summarizes the activities and findings of the project.

Intervention model and components

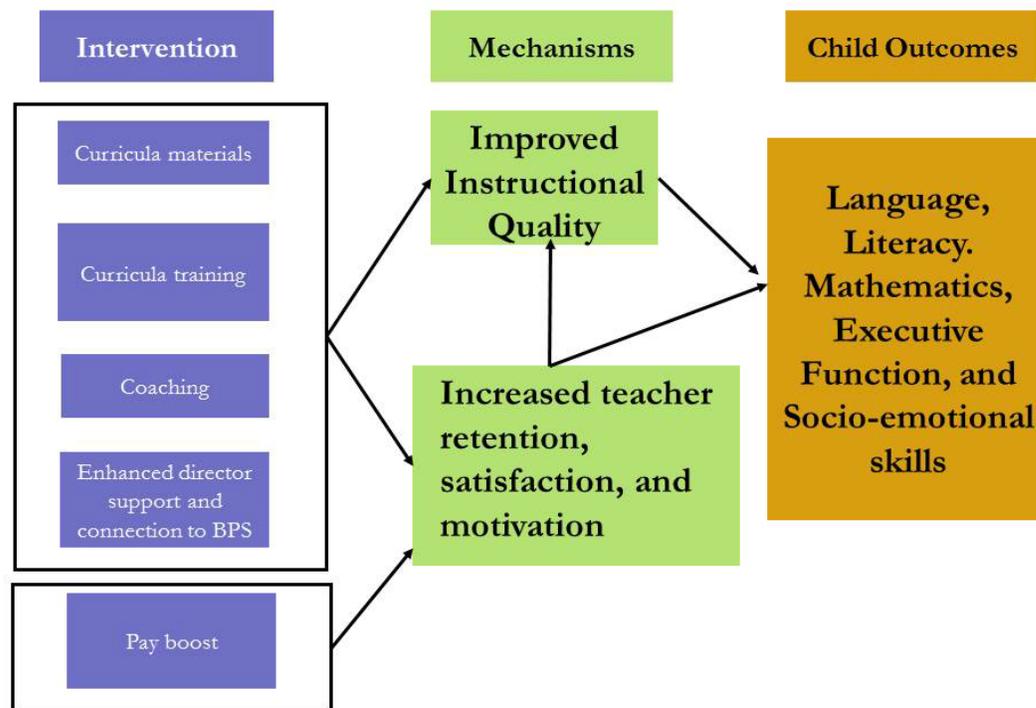
The K1DS theory of change (see Figure 1) hypothesized that child outcomes would improve as classroom instructional quality improved. Sustained improvements in classroom instructional quality were expected to take place over the course of the 2.5-year implementation process.

By participating in this partnership, CBO pre-K teachers received, over a two-year period:

- **Instructional materials and support** to implement the BPS language, literacy, and mathematics curricula (Opening the World of Learning (OWL) and Building Blocks, respectively) in the classroom;
- **Professional development alongside BPS early childhood teachers;**
- **Monthly one-on-one coaching** focused on the language, literacy, and mathematics curricula to translate new knowledge into practice in the classroom; and
- **Supplementation to their salary and benefits** as requested by center directors.

CBO center directors and center-based instructional leaders attended monthly K1DS project meetings, which also served as professional development and information-sharing sessions. These sessions focused on strengthening CBO staff connections to and knowledge of BPS, as well as on strengthening instructional leadership within each center.

Figure 1 The K1DS Theory of Change places emphasis instructional quality and increased teacher retention as levers in improving child outcomes.



Research questions

The evaluation of the CBO K1DS pilot project focused on answering the following research questions:

- Did implementing the Boston model in CBO K1DS classrooms improve instructional quality?
- How did the mathematics, language, and literacy instruction in CBO K1DS classrooms change?
- What were the barriers to implementing the Boston model in CBO K1DS classrooms?
- Did children enrolled in CBO K1DS classrooms show gains in their school readiness skills?

Evaluation

The evaluation design is explained in detail in the full report. Here, we present an overview of our methods. Data on language, literacy, and mathematics instructional quality were collected in each classroom at baseline before the intervention began (January 2013) and at the end of each school year (May 2014 and May 2015). At all time points, we used the Early Language and Literacy Classroom Observation (ELLCO; Smith, Dickinson, Sangeorge, & Anastasopoulos, 2002) to measure language and literacy instructional quality and the Classroom Observation of Early Mathematics Environment and Teaching (COEMET; Sarama & Clements, 2009) to measure mathematics instructional quality. In May 2015, we also used the Classroom Assessment Scoring System (CLASS; Pianta, La Paro, & Hamre, 2008) and we assessed fidelity to curricula in order to understand the degree to which teachers were implementing the K1DS program as it was planned. In the May 2015 measurement period, a sample of BPS K1 ($N=23$) classrooms were rated on the COEMET, ELLCO, and the CLASS.

We also assessed the language, literacy, mathematics, and executive function skills of enrolled children at the start and end of each school year. A trained child assessor, reliable on all measures used, assessed children using six instruments:

- the Peabody Picture Vocabulary Test-III (PPVT-III; Dunn & Dunn, 1997), which measured children's receptive language skills;

- the Woodcock-Johnson Letter-Word subtest (Woodcock, McGrew, & Mather, 2001), which measured children's early reading skills;
- the Applied Problems subtest (Woodcock et al., 2001), which assessed children's early mathematics skills;
- the Forward Digit Span (FDS; Gathercole & Pickering, 2000; Wechsler, 1986), which measured children's working memory;
- the Pencil Tap (Diamond & Taylor, 1996), which measured children's inhibitory control; and
- the Task Orientation Questionnaire (TOQ; Smith-Donald, Raver, Hayes, & Richardson, 2007), which was completed by the assessor after the assessment of each child and was meant to reflect the assessor's perspective on children's attention, impulse control and positive emotion.

In addition to classroom observations and child assessments, we interviewed center directors ($N=10$) and all classroom teachers ($N=34$) three times in the study period (October 2013, June 2014, and June 2015). We also interviewed all BPS and Boston Thrive in 5 staff involved in the CBO K1DS pilot in the fall of 2013 ($N=7$). The BPS-employed CBO K1DS coach was interviewed at three time points (November 2013, June 2014, and June 2015). At the end of each academic year, we collected coaching logs which detailed the topics discussed during coaching sessions as well as follow-up plans. We administered surveys to CBO K1DS teachers in the fall of 2013 and spring of 2014 regarding their satisfaction with the components of the K1DS intervention. Finally, a subset of CBO K1DS parents ($N=48$) completed a survey in the spring of 2015 regarding their motivations for selecting the center in which his or her child was enrolled.

Key Findings

As intended, CBO K1DS reached a different population than BPS K1.

Participating CBO K1DS centers were located in some of the most impoverished neighborhoods in Boston, specifically Boston's Circle of Promise and East Boston, whereas BPS K1 classrooms are located in each neighborhood of Boston. As Table 1 shows, a higher percentage of African American children attended CBO K1DS classrooms than BPS K1. Relatedly, a slightly higher percentage of K1DS students speak English at home than in BPS K1 (67% vs.

50%). Most importantly for the implementation of the OWL and Building Blocks curricula (curricula designed for 4 year olds), not all children in CBO K1DS classrooms were 4-years-old at the beginning of the school year. On average, between 66% (in 2013-2014) and 60% (in 2014-2015) of all children enrolled in CBO K1DS classrooms were 4-years-old. However, it is important to note that classrooms differed widely in the percentage of children who were 4-years-old, ranging from 12%-100%. The classroom in which 12% of the children were 4-years-old at the beginning of the school year had one 4-year-old child in a classroom otherwise composed of 3-year-old children. One recent study found that four-year-old children scored lower on language, literacy, and mathematics kindergarten readiness assessments after attending prekindergarten (in this case, Head Start) classrooms with higher percentages of 3-year-olds (Ansari, Purtell, & Gershoff, 2015). Classroom age composition was not associated with either improvements or decreases in the social-emotional skills of 3- and 4-year-olds.

Table 1 Characteristics of children in CBO K1DS classrooms compared to characteristics of children in the Boston Public Schools K1 program.

Percentage of children who were...	CBO K1DS		BPS K1
	2013-2014	2014-2015	
Latino/a	26%	34%	41%
African American	57%	49%	26%
White	8%	7%	13%
Asian American	6%	6%	11%
Native English Speakers	67%	69%	50%
At least 4-years-old by Sept 1 of school year	66%	60%	100%
Receiving financial assistance to attend prekindergarten (CBOs only)	87%	89%	--
Receiving free/reduced lunch (BPS only)	--	--	69%

Note: CBO KIDS 2013-2014: $N_{\text{children}}=259$, CBO K1DS 2014-2015: $N_{\text{children}}=220$. These demographic characteristics describe all children present in the classrooms in October of the school year (3 and 4 year olds). The financial assistance that children in CBO K1DS classrooms received to attend prekindergarten included MA EEC vouchers and UPK subsidies received by the center. BPS K1 $N_{\text{children}}=2,018$. The BPS sample was from the Preparing to Succeed study which included children who attended the BPS K1 Program in 2008-2009 or 2009-2010 (see Weiland & Yoshikawa, 2013). All children in BPS attend for free.

CBO K1DS teachers differed from BPS K1 teachers in several ways. In Table 2, we summarize teacher characteristics from both CBO K1DS cohorts and a BPS sample. Despite the requirement that all lead teachers in CBO K1DS classrooms had attained at least a BA, only 86%

(2013-2014) to 93% (2014-2015) of CBO K1DS lead teachers had a BA compared to 100% in the BPS K1 classrooms. Because BPS teachers are required to obtain a master's degree within five years of starting a teaching position in the district, 78% of BPS K1 teachers had attained at least a master's degree. This is in comparison to 21% (2013-2014) or 14% (2014-2015) of teachers in CBO K1DS classrooms with master's degrees.

Table 2 Characteristics of Teachers and Classrooms—comparison between CBO K1DS and BPS K1 classrooms.

	CBO K1DS		BPS K1
	2013-2014	2014-2015	
% lead teachers with at least a BA	86%	93%	100%
% teachers with MA	21%	14%	78%
% teachers with 5+ yrs experience	71%	71%	75%

Note: 2013-2014: CBO K1DS $N_{\text{teachers}}=34$; 2014-2015: CBO K1DS $N_{\text{teachers}}=33$; BPS $N_{\text{teachers}}=125$. BPS sample from the Preparing to Succeed study, which included teachers in the BPS K1 program in 2008-2009 or 2009-2010 (see Weiland & Yoshikawa, 2013).

Take-up rates of intervention components was high.

On average, directors attended 70% of the monthly project meetings (range of 58-92%). The majority of the original teachers in participating K1DS classrooms attended the large scale OWL and Building Blocks training sessions (88%, range of 0-100%). Teacher attendance at the @Scale sessions (math-focused professional development sessions offered in the spring of 2015) was high, averaging 92% (range of 0-100%). The BPS K1 coach met with classroom teachers an average of 16 times in the intervention period. This represents roughly 70% of the total opportunities for monthly coaching. Classroom teachers ranged in their uptake of this intervention component with two classrooms receiving as low as 11 coaching sessions (48%) and one classroom receiving as high as 20 coaching sessions (87%).

Roughly \$430,000 was distributed to lead and assistant teachers between January 2013 and June 2015. The amount of money teachers received was determined in part by each center's requested amount at the time of application and teachers' education and experience levels. The median amount supplemented was \$10,584 (range of \$3,790-\$14,324 annually), which represented 27% of teachers' median base salary and wages prior to the

supplementation. The average K1DS teacher made \$13.83 per hour before K1DS and \$22.89 per hour after the K1DS supplementation.

Additional funds were distributed to centers to cover the cost of some curricular materials and the cost of hiring substitute teachers. These substitute teachers were meant to provide coverage for CBO K1DS teachers when they met with the BPS K1 coach or used out-of-classroom planning time.

Intervention components were well-received by K1DS participants, though participants did suggest some changes.

Generally, directors were pleased to have the opportunity to build a professional community with directors from across the Boston area. Directors and teachers universally agreed that the coaching portion of the K1DS program was a major benefit of the program. While some participants requested specific kinds of changes to the coaching model, including more information sharing between the coach and the center director (4 out of 10 directors requested this during the May-June 2014 interview), all described specific ways in which the coaching professionalized the work of teachers and enhanced planning for and implementation of the OWL and Building Blocks curricula.

Feedback on the mathematics professional development sessions was mixed. These were led by BPS instructional research and development staff who specialize in mathematics instruction. Surveys administered directly after the training sessions yielded positive responses to prompts such as *“I will use what I learned in this professional development session this school year”* and *“I learned enough to make this professional development session worth my time.”* The average response to these questions across all eight sessions was equal to 4.5 or above on a 5-point scale (5=strongly agree). However, when interviewed, 80% of teachers reported being frustrated by the pace and content of the sessions. The trainings were planned to increase teacher mathematics knowledge, in accordance with a theory that increased teacher mathematics knowledge is necessary for improving math instruction. However, teachers wanted to learn strategies for implementing Building Blocks within their classrooms. Teachers were less interested in increasing their own knowledge of and comfort with mathematics than they were with gaining concrete and directly applicable knowledge about teaching mathematics

to young children. This disconnect was at the heart of many teachers' disappointment with these sessions.

Finally, in May-June 2015, 6 out of 10 directors indicated that they would like more timely and concrete communication with BPS regarding K1DS project plans as well as plans regarding how the community prekindergarten centers might officially become part of the BPS pipeline into kindergarten.

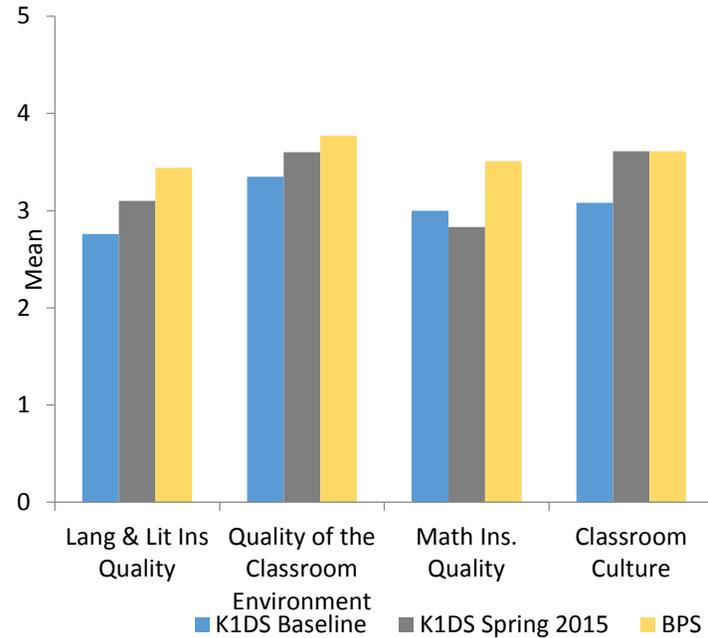
Gaps in classroom quality between CBO K1DS and BPS K1 classrooms shrank over the course of the intervention, but initial progress was fragile.

At the start of the K1DS project, CBO K1DS classrooms lagged behind BPS K1 classrooms in language and literacy and mathematics instructional quality (see Figure 2 Panel A). All of the differences were statistically significant. The standard deviations of these measures is generally small (1 point or less). Accordingly, in Figure 2 Panel B, we translated them into standardized difference (or standard deviation) units; the height of each bar indicates how much lower the quality was in K1DS classrooms compared to the average quality in K1 classrooms before any of the K1DS components had been delivered (baseline) and after 2.5 years of intervention (Spring 2015).. Figure 2 Panel B shows that baseline differences in quality were quite large (e.g. ~ 1 SD or larger) but gaps in quality were reduced by about half or more for three out of four literacy/language and mathematics quality measures. Gaps, however, widened for math instructional quality. Figure 3 compares K1DS and BPS process quality, as measured by the CLASS in spring 2015, both in terms of mean scores (Panel A) and standardized differences (Panel B). Differences were modest for Emotional Support and large and statistically significant for Organization and Instructional Support.

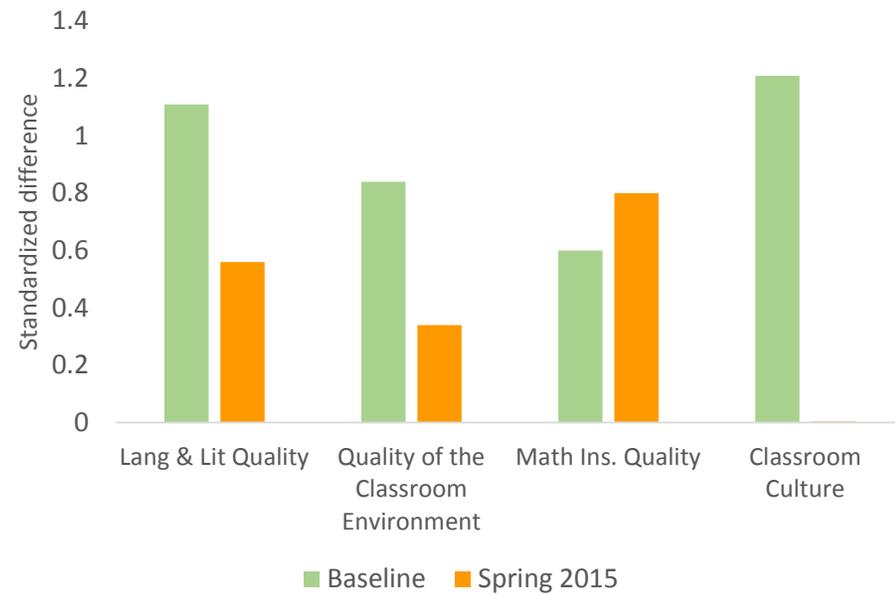
Figure 4 displays K1DS classroom means for the literacy/language and mathematics quality measures across the three time points. For all these measures, there was an increase in quality from baseline to spring 2014 and a decline in quality from spring 2014 to spring 2015. Notably, spring 2015 means are modestly higher than baseline means for three out of the four subscales, meaning some of the initial improvements in literacy/language and mathematics instructional were sustained.

Figure 2 Comparison of CBO and BPK classrooms instructional quality as measured by the ELLCO and COEMET at baseline and after 2.5 years.

Panel A: Means comparison



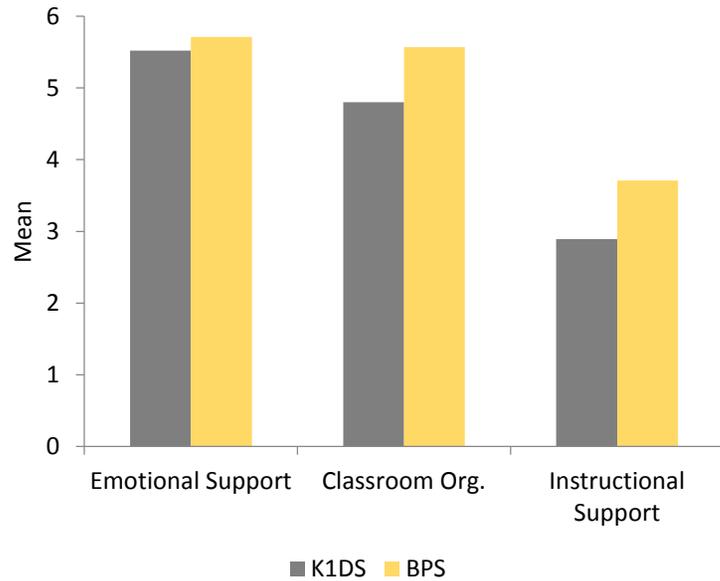
Panel B: Standardized differences comparison



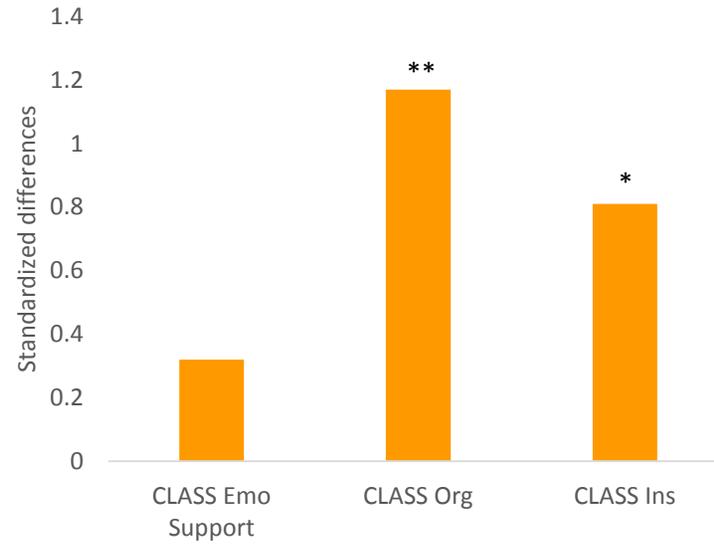
Note: CBO N=13 at baseline and N=10 in Spring 2015. BPS N=23 in Spring 2015. Panel A displays mean scores by time point and auspice only. Panel B displays standardized differences between K1DS and BPS classroom, which were computed by dividing the difference between CBO K1DS and BPS K1 scores by the BPS K1 Spring 2015 standard deviation of the relevant subscale. Differences between CBO K1DS at baseline and BPS K1 classrooms were statistically significantly different for Language and Literacy Instructional Quality and Math Instructional. Quality ($p < .05$). Other mean differences shown in the figure were not statistically significant.

Figure 3 Average classroom instructional quality in K1DS and BPS in Spring 2015, as measured by the CLASS.

Panel A: Means comparison

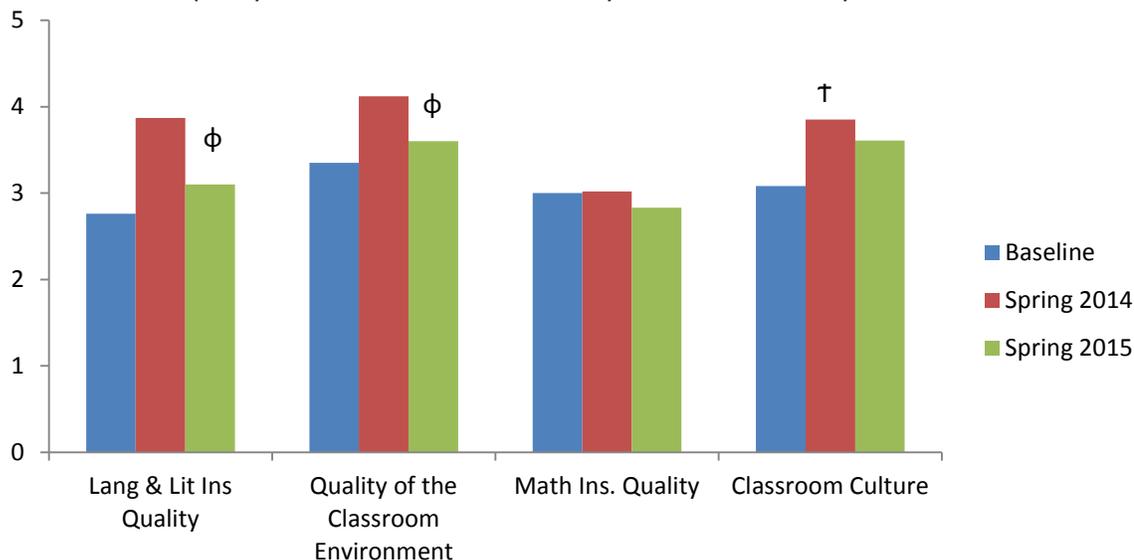


Panel B: Standardized differences comparison



Note: CBO N=10 in Spring 2015. BPS N=23 in Spring 2015. Standardized differences were computed by dividing the difference between CBO K1DS and BPS K1 scores by the BPS K1 Spring 2015 standard deviation of the relevant subscale. Differences between CBO K1DS and BPS classrooms were statistically significantly different in Spring 2015 on CLASS Organizational Support and CLASS Instructional Support ($p < .05$). Other mean differences shown in the figure were not statistically significant.

Figure 4 CBO classroom quality scores at baseline, after 1.5 years, and after 2.5 years



Note: $N=13$ at baseline; $N=14$ in Spring 2014; $N=10$ in Spring 2015. Language and Literacy Instructional Quality and Quality of the Classroom Environment are subscales of the ELLCO. Math Instructional Quality and Classroom Culture are COEMET subscales. ϕ : Spring 2014 scores for both ELLCO subscales were statistically significantly different from scores at baseline and in Spring 2015 ($p<.05$). \dagger : Baseline and Spring 2014 scores on Classroom Culture were statistically significantly different ($p<.05$). Other within-scale means shown in the figure were not statistically significantly different.

Time Spent on Instruction Varied across CBO K1DS Classrooms.

K1DS classrooms have a longer day than BPS, with some running up to ten hours (e.g., from 8:00 am to 6:00 pm) compared to the 6-6.5 hours BPS K1 school day. This is a potential strength of CBOs; more total time presents the opportunity for greater curricula dosage.

Using data from observer notes during fidelity observations in CBO classrooms in spring 2015, we found that approximately 53% of the observed time, or 80 out of 150 minutes, was spent on instruction (range = 43% -73%) and 37% of instructional time was spent in whole group instruction (range = 0% - 60%). Classrooms spent between 9% and 32% of the time on transitions, for an average of 17% of time on transitions.

The overall pattern across CBO K1DS classrooms was that not enough time was devoted to OWL and Building Blocks. In total, core OWL curriculum components require approximately 3 hours per day to implement fully. For Building Blocks, when fully implemented, children are exposed to approximately 180 minutes of Building Blocks per week, through whole group, small

group, center and computer activities.¹ From K1DS classroom schedules, most teachers delivered the core of their Building Blocks and OWL instruction in the morning. Some classrooms did implement some OWL/Building Blocks components in the afternoon – generally a half-hour of additional curriculum time. Adding this 30-minute block to the 80-minute average, CBO K1DS classrooms on average failed to hit the benchmarks for curriculum dosage.

CBO K1DS classrooms varied in fidelity to curricula.

In May 2015, the BPS K1 coach and an observer reliable on the use of the OWL/Building Blocks fidelity measure co-observed CBO K1DS classrooms in order to assess the degree to which teachers were adhering to the expectations of the K1DS project. In Table 3 we summarize these observations and illustrate that 4 of the 14 classrooms had low fidelity of implementation for both curricula; 7 reached a moderate level; and 3 a high level of fidelity.

In Table 4, we provide results of bivariate correlations we conducted in order to understand the relationship between curricular fidelity and classroom quality. We also examined the bivariate correlations between several other factors that project data suggested were facilitators or barriers to higher implementation -- stability in the teaching or leadership team, percent of children under age 4 enrolled in the classroom, and retention of Creative Curriculum—a curriculum used in many CBO preK classrooms prior to K1DS. Due to the small sample size, we stress magnitude of relations and not statistical significance (though *p*-values appear in Table 4).

For both OWL and Building Blocks, higher fidelity was positively associated with higher overall classroom quality, having a stable teaching team, and having a stable director. A higher percentage of children enrolled who were under age 4 and the retention of Creative Curriculum were negatively associated with curricular fidelity.

¹ In BPS's implementation of the two curricula together, Building Blocks is often delivered in conjunction with OWL structures -- e.g., Building Blocks and OWL small groups run simultaneously (teacher leads one, while the assistant leads the other). According, expected total time on OWL and on Building Blocks in BPS is not the same as expected total time when each curriculum is implemented without the other. The time benchmarks presented here are estimates.

Table 3 Classroom-level fidelity to curricula in CBO classrooms, expressed in terms of meeting benchmarks

School	Teacher	OWL benchmark met	Building Blocks Benchmark met	Overall Fidelity benchmark met
1	1	High	High	High
2	2	High	High	High
2	3	High	High	High
3	4	Medium	High	Medium
4	5	Medium	Medium	Medium
5	6	Medium	Medium	Medium
6	7	Medium	Medium	Medium
7	8	Medium	Medium	Medium
8	9	Low	Medium	Medium
4	10	Medium	--	Medium
5	11	Low	Low	Low
9	12	Low	Low	Low
9	13	Low	Low	Low
10	14	Low	Low	Low

Note: 65%+ is high fidelity; >40% is medium fidelity; <40% is low fidelity. Cutoffs were empirically determined for the purposes of our study (e.g., they are not prescribed by the OWL or Building Blocks developers).

Table 4 Results of bivariate correlations conducted to understand the relationship between continuous ratings of fidelity of implementation and classroom quality, both measured in May-June 2015 ($n_{\text{classrooms}}=11$).

	OWL fidelity	p-value	Building Blocks fidelity	p-value
Lang & Lit Ins Quality	0.43	0.21	0.66	~ 0.08
Quality of the Classroom Environment	0.40	0.25	0.62	0.10
Math Ins Quality	0.03	0.94	0.56	0.15
Classroom Culture	0.18	0.63	0.24	0.56
CLASS Emotional Support	0.25	0.49	0.69	~ 0.06
CLASS Classroom Organization	0.53	0.12	0.83	* 0.01
CLASS Instructional Support	0.30	0.41	0.55	0.16
Stable teacher	-0.02	0.94	0.01	0.97
Stable teaching team	0.50	~ 0.07	0.61	* 0.04
Stable director	0.32	0.26	0.67	* 0.02
Percent children under 4-years-old	-0.33	0.25	-0.72	** 0.01
Teachers retained Creative Curriculum	-0.52	~ 0.06	-0.71	** 0.01

Note: Stable teacher refers to the presence throughout the 2.5 implementation period of at least one teacher who entered the project in January 2013.

Nine Barriers to Implementation of the K1DS Theory of Change

Interviews with CBO K1DS directors and teachers as well as with the BPS coach revealed nine barriers to implementation. First, in the period between January 2013 and June 2014, the majority of K1DS teachers reported challenges in planning for two curricula (OWL and Building Blocks) ($N=25$ out of 34 teachers, representing teachers in 10 out of 14 classrooms). These struggles persisted for teachers in 5 of the 14 CBO K1DS classrooms in the 2014-2015 school year ($N=12$ out of 34). Twenty-five teachers (representing 10 out of the 14 classrooms) reported that planning was particularly challenging because the new curricula conflicted with standard practices in the K1DS centers.

Second, implementation was particularly challenging in classrooms where teachers did not stop using part or the entire curriculum in use prior to the K1DS program. This was most evident in the 10 classrooms using Creative Curriculum prior to participating in K1DS. The expectation was that teachers would cease using these curricula, but some teachers persisted using previous curricula because many of the classrooms in these centers were also using Teaching Strategies GOLD as a child assessment and parent communication tool. Teaching Strategies GOLD has a reporting structure that constrains how teachers indicate their plans and learning activities. It also creates a report after teachers have entered data that suggests particular activities and strategies derived from the Creative Curriculum system. Both of these aspects of Teaching Strategies GOLD disrupted the full implementation of the K1DS curricula.

Third, schedules in CBOs posed problems for children's exposure to intentional instruction. Teachers in 12 of the 14 classrooms struggled to implement more than 2 hours of planned, intentional instruction in the 10-hour day. This seemed due in part to the fact that 1) children arrived to the classroom in a staggered fashion, 2) the lead teacher only worked a portion of the available instructional time by either working the early shift—starting at 7:30 am and leaving at 3:00 pm or the late shift—starting at 10:00 am and leaving at 6:00 pm, and 3) CBO K1DS centers scheduled nap times and two meals into their days. Additionally, some teachers indicated that they took breaks from the K1DS curricula especially around holidays or during times when many children were out for vacation. Overall, the more flexible expectations

in CBOs around the start and end of the children's school day and the number of school days attended appears to have led to reduced curricular dosage and adherence.

Fourth, lack of common planning time for K1DS teachers has also presented a barrier to change. Only one of the 10 K1DS centers provided reliable coverage for consistent out-of-classroom planning time. For teachers in the other nine centers, this meant that planning time often occurred during the children's nap time. Additionally, teachers from centers without regular planning time reported taking turns to plan on weekends or after work and communicating with the other teaching team members through notes or brief conversations during the children's free choice time. Twenty-four (spring 2014; 71%) to 27 (spring 2015; 79%) teachers reported on this challenge.

Fifth, nine out of 10 center directors reported that finding and paying for predictable coverage for teachers was an on-going issue in both the 2013-2014 and 2014-2015 school years. This was an issue even though centers that participated in the CBO K1DS program received funds for substitute teachers. This issue interfered with centers' ability to (1) schedule coaching sessions, (2) provide predictable, frequent out-of-classroom planning time for teaching teams, (3) and institute a regular accountability system that included classroom observations and teacher-director debriefs. Relatedly, center directors reported challenges in recruiting qualified teachers (particularly those with a bachelor's degree) in light of the compensation discrepancy between average CBO pay and BPS pay. This was at the heart of two centers' struggle to replace CBO K1DS teachers who had left prior to the 2014-2015 school year. In one of these centers, few qualified candidates applied. In the other center, a replacement teacher left mid-year after being offered a higher-paying job.

Sixth, five out of 10 centers had only one full-time administrator. Directors in these centers, plus three additional directors (8 out of 10), reported not having sufficient time to regularly observe classrooms and provide instructional support and professional development focused on improving classroom quality. Relatedly, only two center directors reporting having a systematized accountability system which provided clear goals and expectations for teacher performance. This issue was not remedied in the second year of the program.

Seventh, four teachers in two of the 14 classrooms reported struggling more with challenging child behaviors in 2014-2015 compared to the prior year. These teachers did not feel adequately supported in meeting the needs of these children. These teachers felt that managing these challenging behaviors undermined fidelity to the K1 model.

Eighth, most CBO K1DS classrooms were mixed-age despite the program's original intent. Twelve classrooms enrolled both 3 and 4-year-old children and the average percentage of 3-year-old children in these classrooms was 66% (2013-2014) to 60% (2014-2015) (range of 12% to 100% in both years). This is important because the OWL and Building Blocks curricula were developed for four year olds. The K1DS coach reported that teachers struggled to differentiate the curricula adequately for both age groups, which undermined curricular implementation and instructional quality.

Finally, gains in quality may have been hard to sustain through the end of the project due to a relatively high rate of turnover (50% for teachers from baseline to spring 2015). This turnover reduced the capacity of classrooms to implement the K1DS program components because: 1) many centers struggled to replace teachers, 2) once hired, new staff did not receive formal training on the curricula used in K1DS, and 3) collaboration within teaching teams was destabilized by the turnover.

Children attending K1DS classrooms showed gains in their school readiness skills.

In this exploratory pilot study, child assessment data were meant to ground our understanding of child learning in community-based K1DS classrooms and were not meant to provide information about the success (or lack thereof) of the CBO K1DS project. This is particularly the case because we lacked a matched comparison group from the district and were not able to randomize children to the CBO K1DS classrooms. Furthermore, the mixed-age groups present in 12 of the 14 classrooms means that the majority of the CBO K1DS children experienced OWL and Building Blocks implementation in a way that was potentially different from children in classrooms in which only 4-year-olds were in attendance.

Overall, children's performance on direct assessments of their receptive language, literacy, mathematics, and impulse control improved from the start of their prekindergarten year to the end of that same year. In Table 5, we present a summary of student performance on

child assessments from CBO K1DS classrooms. In this table, we have averaged child assessment scores across both academic years. Cohort 1 contained 100 4-year-old children who attended CBO K1DS classrooms in the 2013—2014 school year. Cohort 2 contained 63 4-year-old children who attended CBO K1DS classrooms in the 2014-2015 school year. Given the small sample size of each cohort and the relatively stable level of classroom quality across both years, we averaged the child scores across both cohorts before comparing the performance of children from CBO K1DS classrooms with the performance of children from BPS K1 classrooms. In column 4 we list the effect sizes of a quasi-experimental study of the impact of attending a BPS K1 classroom. The effect sizes represent the standardized difference in scores between children who attended K1 and those who did not. For details about this study, please see: (Weiland & Yoshikawa, 2013). In columns 5, 6, and 7 we describe the performance of a subset of children who attended BPS K1 classrooms in the 2009-2010 school year. These children were only assessed on a subset of the instruments used in this study; therefore, there are blank rows in these columns. Where possible, we compare the scores of CBO K1DS children to children represented in columns 5-7.

It is important to remember that this study did not have an experimental design and that there are many unobserved ways in which the CBO K1DS children differed from the As the last column of this table illustrates, children who attended CBO K1DS classrooms scored lower than children who attended BPS K1 classrooms (represented by a positive value in this column) in all but one domain. CBO K1DS children performed as well or better, on average, on an assessment of their receptive language. We provide further details about child assessment performance in the Appendix.

Table 5 Comparisons of CBO K1DS and BPS K1 child performance on a range of child assessments.

	<i>CBO K1DS Cohorts 1 & 2</i>			<i>BPS impacts</i>	<i>BPS 2009-2010</i>		<i>CBO-BPS differences</i>	
	<i>Fall</i>	<i>Spring</i>	<i>Stand</i>	<i>ES</i>	<i>Fall</i>	<i>Spring</i>	<i>Stand</i>	
	<i>Mean</i>	<i>Mean</i>	<i>difs</i>	<i>(SD)</i>	<i>Mean</i>	<i>Mean</i>	<i>difs</i>	
	<i>(SD)</i>	<i>(SD)</i>			<i>(SD)</i>	<i>(SD)</i>		
Receptive language	88.13 (16.53)	95.98 (13.15)	0.45	0.44	88.19 (17.63)	94.45 (17.89)	0.36	-0.09
Early reading	9.13 (5.21)	10.67 (7.41)	0.28	0.62	--	--	--	0.34
Early math	10.76 (5.34)	12.36 (7.29)	0.34	0.59	--	--	--	0.25
Working memory	4.07 (0.94)	4.41 (0.91)	0.26	0.24	3.86 (1.31)	4.46 (1.18)	0.46	0.20
Inhibitory control	7.69 (5.33)	10.42 (5.02)	0.42	0.21	8.69 (6.47)	12.94 (4.56)	0.66	0.24

Note: CBO KIDS 2013-2014: $N_{\text{children}}=100$. CBO KIDS 2013-2014: $N_{\text{children}}=63$. BPS K1 $N_{\text{children}}=2,018$. Standardized differences for the receptive language, working memory, and inhibitory control measures are reported in terms of the BPS 2009-2010 Preparing to Succeed follow-up study. Standardized differences for all other measures are reported in terms of the Preparing to Succeed control group standard deviations for each measure. The BPS sample was from the Preparing to Succeed study which included children who attended the BPS K1 Program in 2008-2009 or 2009-2010 (see Weiland & Yoshikawa, 2013). Effect sizes are listed as were reported in (Weiland & Yoshikawa, 2013) and are also reported in terms of the control group standard deviations. CBO-BPS differences compare the ES of the CBO classrooms on average to the Preparing to Succeed 2009-2010 ES.

CBO classrooms offer unique strengths.

CBO K1DS classrooms offered several notable strengths as a preschool delivery setting. For example, teachers and students in most centers had family-style meals and some teachers took advantage of this time to engage children in rich conversations characterized by multiple turn-taking and that drew on student interests and experiences. These kinds of conversations build important child language skills and help prepare children for elementary school. In addition, because most programs do not provide transportation, a caregiver for each child was usually in the classroom once or twice a day. Conversations between teachers and caregivers at drop-off and pickup facilitated closer family-school relationships and family engagement. BPS K1 teachers, conversely, generally do not eat lunch with children and many children arrive at school by bus. Finally, the CBO K1DS school day was longer than the BPS K1 school day by an average of 3.5 hours. This means that, in theory though it was not found in this study, CBO K1DS teachers had more time to deliver the language, literacy, and mathematics curricula than did BPS K1 teachers.

Recommendations

Given the findings of the CBO K1DS project, summarized here and unpacked in detail in the full report, we offer the following recommendations as BPS moves forward with this scale-up and scale-out process.

Policy

- **Classroom quality, particularly instructional quality can be systematically improved in community-based prekindergarten programs.**
 - **Positive change in instructional quality was fragile because classroom quality in CBO K1DS classrooms was sensitive to teacher and leader turn-over, the presence of children with challenging behaviors, among other factors.**
 - **Sustainable growth in classroom quality takes time and investment. In the CBO K1DS program, programs benefitted most from targeted coaching support.**
- **A CBO center's capacity for offering a range of systematic professional supports for teachers and leaders makes a difference.**
 - **Centers with the infrastructure to provide stable instructional leadership, predictable out-of-classroom planning time, an accountability system for classroom relational and instructional practices experienced the highest, sustained levels of classroom quality.**
 - **A center's capacity varied by organization size and structure.**
- **Balance support with accountability**

- Provide supports to centers but also put in place a clear accountability system for implementation and performance (Duncan & Murnane, 2013). Centers generally did not have in place strong teacher and director review processes or a strong culture of accountability. Link program participation to meeting a set of clear performance standards.

Practice and Implementation

- Provide more interaction between K1DS teachers and master implementers of Opening the World of Learning and Building Blocks.
 - Classroom visits or video clips offer opportunities to understand the implications of high levels of implementation of the curricula.
- Make clear that the curricula are the critical component of delivering the BPS K1 model and not a supplement or add-on.
 - Classrooms should suspend the bulk of activities associated with former practice or formerly used curricula
 - In particular, integrate clear implementation fidelity expectations into training, coaching, and program quality monitoring. Link program participation to meeting these expectations.
- Focus implementation efforts on classrooms which enroll 4-year-old children only.
 - When this is not attainable, restrict the number of three year olds enrolled and provide extra supports for differentiating the curricula.
- Make training and coaching adaptive, in anticipation of higher teacher turnover
 - Elevated rates of teacher turnover relative to BPS are likely to continue in any CBO expansion effort, due to factors such as low compensation among others. Revise the current training model to adapt to this reality, including formalizing how to bring new teachers up to speed via videotapes of training and additional resources for coaching.
- Start small because piloting:
 - Highlights unanticipated barriers to implementation fidelity
 - Allows program designers to adapt components to suit the capacity of target centers and classrooms
 - Provides an opportunity for participants in the program to provide input on components, pacing, and capacity
 - Reveals baseline capacity needs for any center or classroom wishing to participate in the program
- Support the inclusion of regular, shared planning time for all teaching teams on a predictable schedule.
 - Build a center-wide schedule that allows coverage for classrooms during rest time so that at least two teaching team members are available to plan outside the classroom. Some centers have attained this by staffing all classrooms with three teachers.
 - Relatedly, ensure that qualified substitute teachers are available and that the core teaching teams in K1DS classrooms are not regularly responsible for covering teacher absences in other classrooms.

- **Take steps to support stable teaching teams and leadership.**
 - **Living wages and benefits packages may reduce turnover**
 - **Professionalization of the early childhood workforce may be attained through better instructional and classroom management support by center leadership.**
 - **Relatedly, support staff may help reduce center leadership administrative duties freeing leaders to provide professional support to teachers.**
 - **Create a collaborative space for community-based K1DS center leadership that offers specific, actionable support for data-driven instructional leadership and task management.**

Research

- **Design an evaluation plan that is both summative and formative**
 - **Offer real-time, actionable findings to key stakeholders so that the program model can adapt to the realities discovered during implementation.**
 - **Mixed methods approaches allow all participants a chance to share their unique perspective on the program.**
- **Researcher-practitioner partnerships reduce “translation” gaps by forcing stakeholders to develop and use a shared understanding of program goals, practitioner challenges, and the trade-offs involved in educational change.**

This page was intentionally left blank.

Table of Contents

Preface	ii
Glossary.....	iii
Executive Summary.....	4
Background	4
Intervention model and components.....	4
Research questions.....	6
Evaluation	6
Key Findings	7
As intended, CBO K1DS reached a different population than BPS K1.....	7
Take-up rates of intervention components was high.....	9
Intervention components were well-received by K1DS participants, though participants did suggest some changes.....	10
Gaps in classroom quality between CBO K1DS and BPS K1 classrooms shrank over the course of the intervention, but initial progress was fragile.	11
Time Spent on Instruction Varied across CBO K1DS Classrooms.	14
CBO K1DS classrooms varied in fidelity to curricula.....	15
Nine Barriers to Implementation of the K1DS Theory of Change	17
CBO classrooms offer unique strengths.	22
Recommendations	22
Introduction	28
Boston Prekindergarten Background	28
Boston pursues a mixed-delivery approach	31
Research questions.....	32
The K1DS Pilot Program: Can BPS K1 work outside the public schools?.....	33
CBO site requirements and application process	34
Characteristics of the CBO K1DS Participants.....	35
The Intervention Components.....	40
Curricula materials.....	40
Curricula training	41
Coaching	41
Monthly Director’s Meetings.....	42
Compensation Supplementation.....	43
Participation in K1DS Intervention Components.....	44
Methods for Collecting Data and Evaluating Progress	45

Measures	47
Analytic Approach	55
Findings	56
Teacher and Director Readiness to Change	56
Intervention components were well-received by K1DS participants	57
Classroom Quality	62
Nine Barriers to Implementation	69
Teacher Retention	72
CBO classrooms offer unique strengths	73
Putting it all together: How do Fidelity, Quality, and Child Performance relate to one another?	73
Conclusions	75
Recommendations	76
References	79
Appendix	83

Introduction

For years, evidence has been accumulating across disciplines as diverse as economics and neuroscience that a child's earliest experiences form a lasting foundation upon which all subsequent learning and development grows (see, for example: Center on the Developing Child, 2010). Decades of research have found that high-quality child care and early education settings have the potential to offset the kind of adversity that young children living in poverty are likely to face (Yoshikawa, et al., 2013).

It was in this context that the current system of publicly funded, school-based prekindergarten in Boston arose. In 2005, then-Mayor Thomas Menino called on the Boston Public Schools (BPS) to offer universal prekindergarten (called K1 in the district and referred to as such throughout the remainder of this report). The city, highly segregated along socioeconomic and racial/ethnic lines, was investing more in its poorest neighborhoods. Building school-based prekindergarten programs was one of the ways that city leaders were working to improve resources and infrastructure in the city writ large. Leading to the current study's focus on the K1DS expansion, the number of slots available then and now was not sufficient to meet parent demand.

Boston Prekindergarten Background

From the start, the BPS K1 program had strong structural supports. Teachers were paid on the same scale and were subject to the same educational requirements as K-12 teachers (e.g., all BPS teachers are required to earn a master's degree within 5 years of being hired). K1 classrooms were located in public schools; therefore, K1 teachers had access to on-site supports for special education and dual language learner students. The program was also open to all families, regardless of income. These features are unusual in public prekindergarten programs nationally, which are more often characterized by lower pay and educational requirements for prekindergarten teachers and by mixed-delivery systems in which some or all classrooms are located outside public school buildings (Barnett, Carolan, Squires, Brown, & Horowitz, 2015).

Despite these strong investments, early signs suggested process quality was lacking. In a quality study conducted in 2006, 70% of BPS prekindergarten classrooms were found to have

inadequate quality (Jan, 2007). At that time, classroom teachers were using a range of curricula. Furthermore, paraprofessionals (assistant teachers) were not present in all prekindergarten classrooms. In light of these findings and under the guidance of Dr. Jason Sachs, BPS undertook a dramatic transformation of its K1 program. Prior to joining BPS as head the Department of Early Childhood Education, Dr. Sachs' work focused on classroom quality. Through this work, he came to view formal curriculum as central to the establishment and maintenance of the level of classroom quality so often linked to favorable child outcomes in research literature (J. Sachs, personal communication, July 10, 2014). Therefore, the model that Dr. Sachs and his team developed for the BPS K1 program had a strong focus on research-based, systematically deployed curricula that were aligned to a shared set of learning standards.

Specifically, Dr. Sachs and his team selected two research-based curricula to be used in tandem across all district K1 classrooms —Opening the World of Learning (OWL) and Building Blocks: Foundations for Mathematical Thinking (Building Blocks). The OWL curriculum targets children's early language and literacy skills and includes a social-skills component embedded in each unit, in which teachers discuss social-emotional issues with children and integrate emotion-related vocabulary words (Schickedanz & Dickinson, 2005). The Building Blocks curriculum targets early mathematics skills, particularly 1) number and simple arithmetic, and 2) geometry, measurement, and spatial sense. Three mathematical themes – patterns, data, and sorting and sequencing – are woven into these two main areas. In addition, many activities are intentionally child-directed, with children making up their own problems or creating their own geometric designs (Clements & Sarama, 2007a). Its pedagogical approach has a heavy focus on language, as it requires children to explain their mathematical reasoning verbally.

Studies of OWL and Building Blocks have shown positive effects on directly targeted children's outcomes (Ashe, Reed, Dickinson, Morse, & Wilson, 2009; Clements & Sarama, 2007b; Clements et al., 2011).² Other studies have also shown that Building Blocks has cross-

² As detailed in Weiland and Yoshikawa (2013, p. 2116), "the evidence base for Building Blocks is stronger than that for the OWL. Children in eight programs that implemented OWL showed consistently positive effects in studies that used pre-post designs with no control group (Wilson, Morse, & Dickinson, 2009). But a recent randomized controlled trial in Head Start centers (Dickinson, et al., 2011; Dickinson, Freiberg, & Barnes, 2011) found no impacts of OWL on children's language and literacy outcomes at the end of preschool, and some negative effects at the end of kindergarten and the end of first grade. However, these latter results are somewhat difficult to interpret, as the

domain effects on children's executive function skills and on four measures of oral language including the ability to recall key words, use of complex utterances, willingness to reproduce narratives independently, and inferential reasoning on practical content (Sarama, Lange, Clements, & Wolfe, 2012).

Importantly, these curricula were not used in isolation. Their implementation was supported by curriculum-specific training and weekly to bi-weekly on-site support from an experienced early childhood coach trained in both curricula. In the first year of implementation, teachers were offered two days of curricular training in Building Blocks and five days in OWL. During the school year, teachers were offered four days of training in Building Blocks and two days of training in OWL. In the second year of implementation, all teachers new to the K1 program were offered five days of curricular training before the start of the school year and six days of training during the school year. Coaching sessions were tailored to address the individual needs of each teacher in implementing the curricula and managing the classroom.

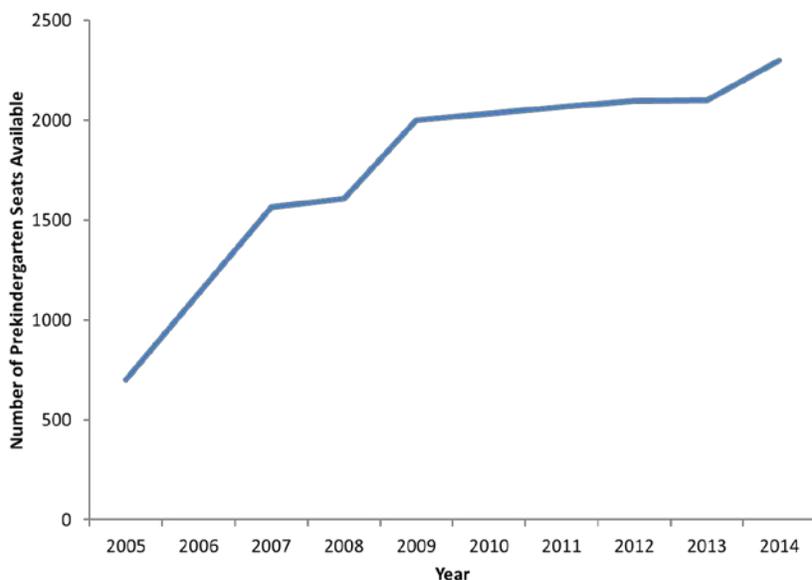
In the 2009-2010 school year, a quasi-experimental evaluation of the BPS K1 model was conducted to test the efficacy of the model in improving kindergarten readiness for participating children. The study found that K1 had moderate to large impacts on the developmental domains directly targeted by the intervention (language, literacy, and mathematics). There were also modest impacts on children's socio-emotional skills (specifically, on emotion recognition) and modest impacts on all three components of children's executive function skills, a domain not directly targeted by the intervention (Weiland & Yoshikawa, 2013). Of particular interest, average instructional quality was high relative to classrooms in other large studies of prekindergarten instructional quality (Weiland, Ulvestad, Sachs, & Yoshikawa, 2013). Impacts on children were especially pronounced for children from low-income families, dual language learners (Weiland & Yoshikawa, 2013), and children with special needs (Weiland, 2016a).

fidelity of implementation in the treatment groups was relatively low and control classrooms had partially implemented the OWL. Teachers were also on average better educated in the eight programs that showed positive effects than in the RCT (65% vs. 17% with a BA, respectively)."

Boston pursues a mixed-delivery approach

After a rapid 8-year expansion (see Figure 1), BPS pursued the mixed-delivery system that is the focus of the present evaluation for three primary reasons. First, parent demand for the program outstripped supply³ and physical space in public schools for more prekindergarten seats in high-need geographic regions was limited. Second, there was a perception that the BPS K1 model did not meet the needs of a subset of Boston’s working families, particularly families who needed full-day, year-round education and care services for their children and who struggled with logistics of utilizing the BPS K1 program (9 months of 8:30-3:30 pm education and care services; M. High-McKinnon, personal communication, September 24, 2014; J. Tewksbury & A. Shapiro, personal communication, October 30, 2013). Third, there was a statewide emphasis on local multiple-auspices system of early care and education.

Figure 1 Number of BPS K1 seats available from 2005-2014



Accordingly, in 2011-2012, Dr. Sachs and his team of in-house and community partners sought alternatives outside of the school district for expanding the publicly funded

³ While it is difficult to determine precisely how many parents preferred a BPS K1 spot for their child but were unable to obtain one, between 20%-35% of Boston families in 2013-2014 year were not able to place their children in BPS K1 classrooms despite a desire to do so. This range was calculated by comparing the size of the Kindergarten program to the K1 program as well as drawn from an interview with Dr. Sachs regarding the size of the K1 waiting list (J. Sachs, personal communication, July 10, 2014). In 2009, approximately 43% of the children who enrolled in BPS kindergarten had attended BPS K1 (Weiland & Yoshikawa, 2013).

prekindergarten program. Boston, like many other urban centers, has a large and diverse set of community-based organizations (CBOs) that provide early care and education services. Organizational structures and curricular approaches differ a great deal across CBO programs. These differences motivated the careful piloting work that is the subject of this evaluation. That is, BPS and its partners recognized that their model would need to be adjusted to be successful in a context as variable as the community context was.

Curricular choice and professional development were two of the key differences recognized at the outset of this project between BPS K1 and the CBO prekindergarten classrooms. As explained above, BPS K1 uses domain-specific, research-based curricula supported by training and regular on-site coaching. In contrast, many community-based prekindergarten classrooms use global curricula such as Creative Curriculum, which tend to lack the specified scope and sequence that characterize some domain-specific curricula and are not supported by rigorous research (What Works Clearinghouse, 2013). Global curricula are the most widely used nationally (Weiland, 2016b). Furthermore, most CBO prekindergarten teachers do not have consistent access to a systematic form of professional development.

Two additional differences were teacher compensation and accountability. BPS K1 teachers are held to the same accountability standards as district kindergarten through grade 12 teachers and are paid on the same salary scale. Monitoring and evaluation processes in community-based prekindergarten centers are not standardized across settings and oftentimes center directors act more as business administrators than instructional leaders (M. High-McKinnon, personal communication, November 6, 2013). Additionally, community-based prekindergarten teachers are among the worst paid professional group in the United States (U.S. Bureau of Labor Statistics, 2012; Whitebook, Phillips, & Howes, 2014). These differences were also taken into account when the K1 in diverse settings (K1DS) program was designed.

Research questions

The following four research questions arose from the background and research from which the K1DS project arose:

1. Did implementing the Boston model in CBO K1DS classrooms improve instructional quality?

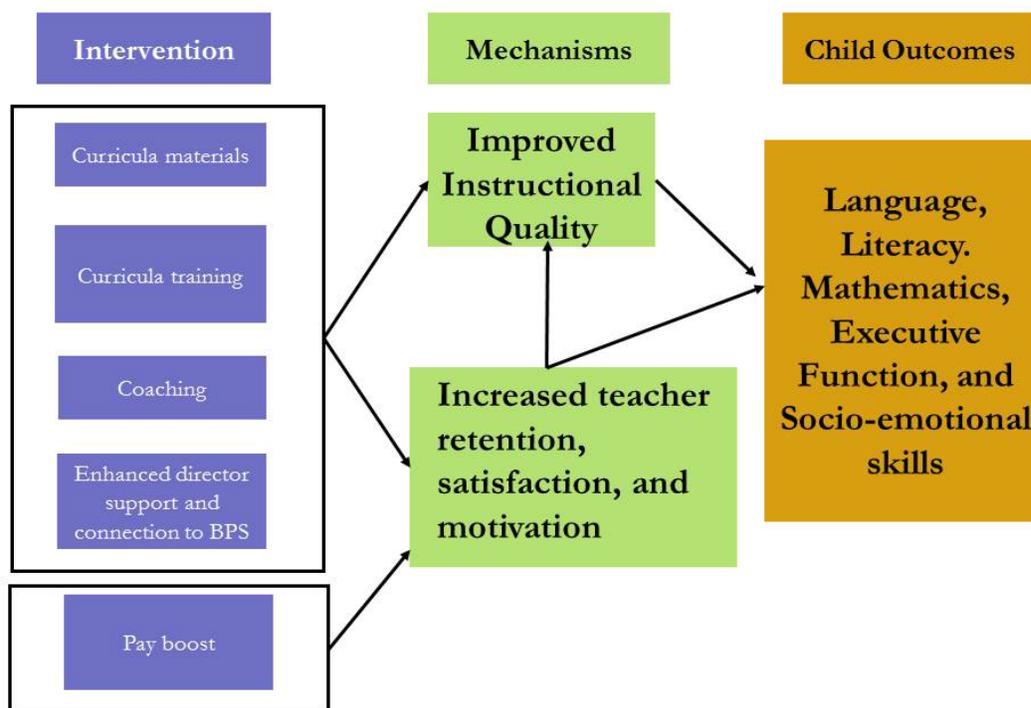
2. How did the mathematics, language, and literacy instruction in CBO K1DS classrooms change?
3. What were the barriers to implementing the Boston model in CBO K1DS classrooms?
4. Did children enrolled in CBO K1DS classrooms show gains in their school readiness skills?

The K1DS Pilot Program: Can BPS K1 work outside the public schools?

As the theory of change in Figure 2 shows, the ultimate goal of the K1DS partnership was to increase child outcomes across domains. The K1DS theory of change posited that classroom instructional quality would increase if the following activities occurred. First, CBO K1DS teachers received materials and professional development centered on the OWL and Building Blocks curricula. Second, a BPS K1 coach first observed and then met with teaching teams on a monthly or bimonthly basis to provide individualized feedback and support. The BPS K1 coach also provided modeling of key classroom processes where necessary. Third, teachers received a salary and benefits supplement. The amount of this supplement was based on requests made in each center’s application to the K1DS program.⁴ Fourth, center directors and center-based instructional leaders attended monthly K1DS project meetings which also served as professional development and information-sharing sessions.

⁴ For a subset of centers—those that entered in Round 2 under the auspice of Boston Thrive in 5, the supplement was retroactive—money was distributed only after the first year of implementation.

Figure 2 The K1DS Theory of Change places emphasis instructional quality and increased teacher retention as levers in improving child outcomes.



CBO site requirements and application process

In December 2011, BPS released a call for applications for CBOs to apply to the K1DS program. There were eight primary requirements that CBO centers had to meet in order to apply for the program. BPS put these criteria in place as they anticipated only teachers in structurally strong centers would be successful in implementing the K1 model. Centers must have been:

1. Licensed by the MA Dept. of Early Education and Care,
2. Located in high-poverty neighborhoods in Boston,
3. Accredited or willing to pursue accreditation with the National Associations for the Education of Young Children (NAEYC), and
4. Operating on a full-day, year-round basis.

Additionally, classrooms must have had:

5. At least 80% of enrolled children residing in Boston,
6. A lead teacher with at least a bachelor's degree, 4-6 courses in early childhood education and three years of EC teaching experience,

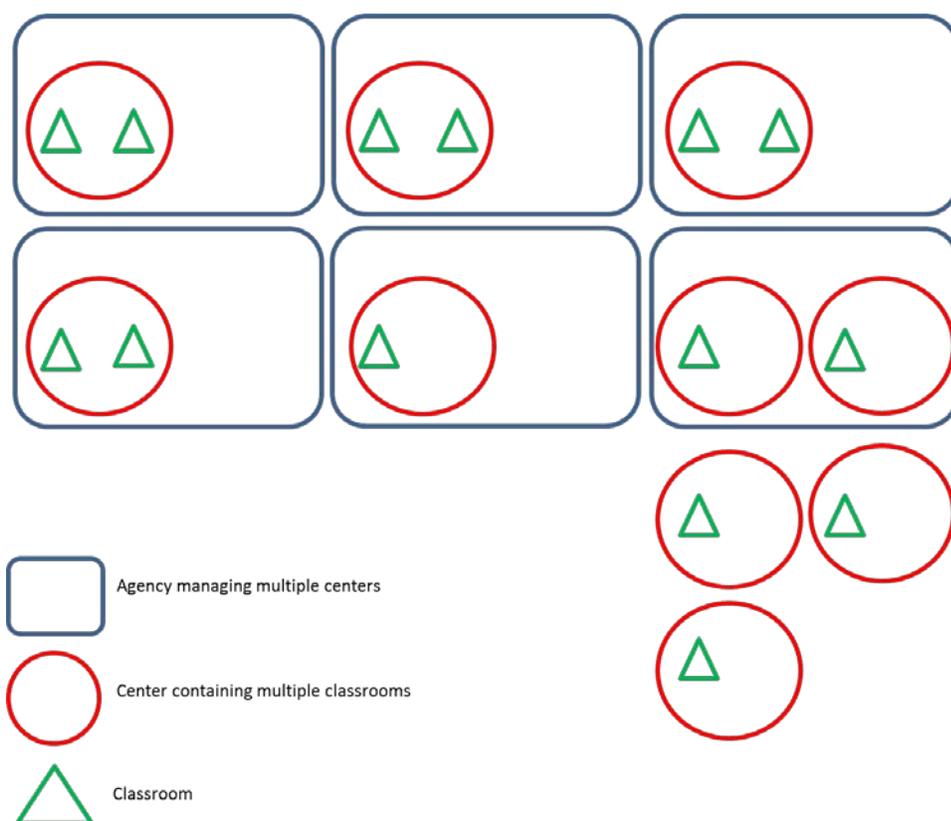
7. An assistant teacher with at least an associate's degree in early childhood or a CDA and one year of teaching experience, and
8. 1:10 teacher-student ratio.

Eleven programs applied to BPS's request for proposals. Of these, seven centers were accepted into the program by BPS in the first round of awards. In a second round, Boston Thrive in 5, a local non-profit agency devoted to improving the learning and school readiness of Boston's most vulnerable children, provided funding to add in three additional applicant centers that had not met at least one of the 8 requirements. Specifically, two of the centers did not have a lead teacher with a BA degree and one center proposed a classroom with a slightly higher child: teacher ratio than requested. In total, 14 CBO prekindergarten classrooms, nested in 10 centers and 6 larger organizations, participated in the January 2013-June 2015 K1DS pilot program.

Characteristics of the CBO K1DS Participants

Organization and Location of the CBO K1DS Centers. Figure 3 illustrates the organization of the 14 CBO classrooms in centers and, where applicable, sponsoring agencies. Participating CBOs represented a wide range of auspice type. The larger agencies involved were all non-profit and one was religiously affiliated. All but one of the agencies provided other services to low-income families in addition to early care and education programs. Three centers were stand-alone and were not affiliated with a larger agency.

Figure 3 Organizational structure of participating CBO K1DS centers.



The neighborhoods served by these centers were among the poorest in Boston. Figure 4 shows a map of Boston with the percentage of children living in families at or below the federal poverty line in between 2005-2009. With two exceptions, K1DS centers are located in the darker purple areas, indicating the highest child poverty rates in Boston. The one center in the Back Bay neighborhood—one of the more affluent areas of Boston—specifically targeted low-income families in Back Bay and the surrounding areas. Table 1 summarizes a short set of characteristics of the seven neighborhoods in which the CBO K1DS classrooms were located.

Figure 4 Percent of families with children living in poverty in Boston neighborhoods.

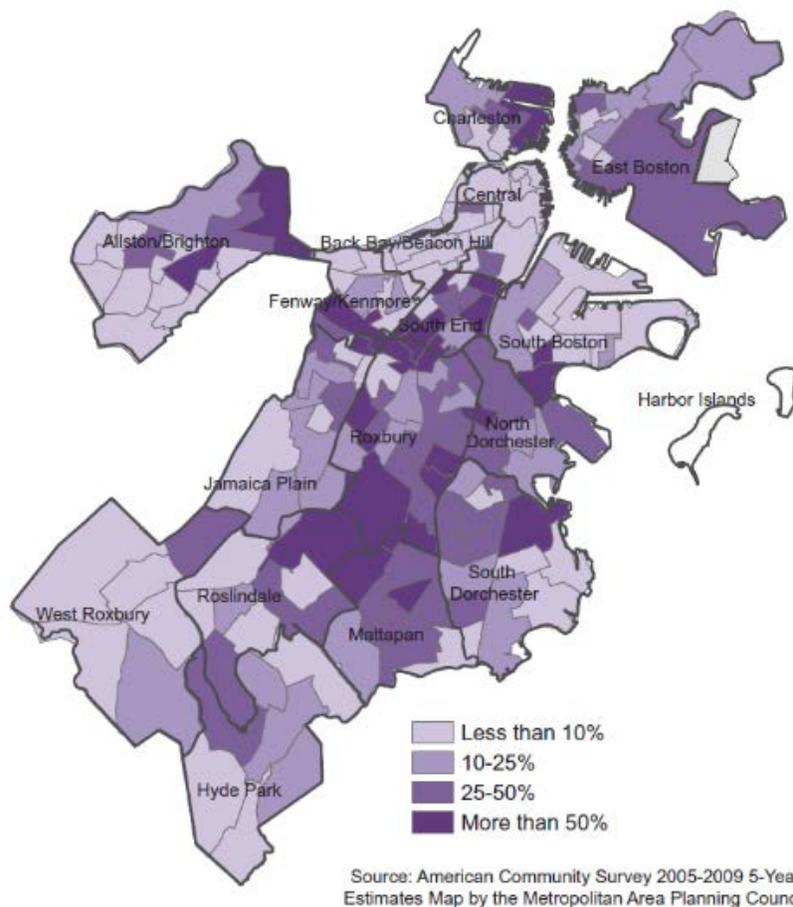


Table 1 Family income characteristics of the seven neighborhoods in which CBO K1DS centers were located.

	N CBO K1DS Centers	Median family income	% Households receiving Food Stamps/SNAP	Poverty rate among all households	Poverty rate among households with children
Back Bay	1	\$ 182,190	3%	13%	2%
Dorchester	3	\$ 48,254	24%	24%	22%
East Boston	1	\$ 47,198	18%	14%	20%
Jamaica Plain	1	\$ 84,045	13%	21%	15%
Mission Hill	1	\$ 36,237	21%	39%	22%
Roxbury	2	\$ 32,432	36%	35%	32%
South End	1	\$ 72,063	17%	13%	17%
Boston		\$ 61, 109			23%

Source: 2007-2011 American Community Survey, BRA Research Division Analysis

All participating centers were NAEYC accredited and offered year-round, full-day services for children between the ages of 2 months and 5 years. Centers differed in their internal organizational structure. While all K1DS centers employed a center director, only three employed an administrator whose main responsibility was instructional support.

CBO K1DS teachers differed from BPS K1 teachers in several ways. Table 2 contains a summary of teacher characteristics from both CBO K1DS cohorts and a BPS sample. Despite the requirement that all lead teachers in CBO K1DS classrooms be staffed with a lead who had attained at least a BA, only 86% (2013-2014) to 93% (2014-2015) of lead teachers had a BA compared to 100% in the BPS K1 classrooms. Because BPS teachers are required to obtain a master's degree within five years of starting a teaching position in the district, 78% of BPS K1 teachers had attained at least a master's degree. This is in comparison to 21% (2013-2014) or 14% (2014-2015) of teachers in CBO K1DS classrooms with master's degrees.

Table 2 Characteristics of Teachers and Classrooms—comparison between CBO K1DS and BPS K1 classrooms.

	CBO K1DS		BPS K1
	2013-2014	2014-2015	
% lead teachers with at least a BA	86%	93%	100%
% teachers with MA	21%	14%	78%
% teachers with 5+ yrs experience	71%	71%	75%

Note: CBO $N_{\text{teachers}}=35$; BPS $N_{\text{teachers}}=125$; CBO $N_{\text{children}}=259$. All percentages are averages across both school years; BPS $N_{\text{children}}=2,018$. BPS sample from the Preparing to Succeed study, which included children who attended the BPS K1 program in 2008-2009 or 2009-2010 (see Weiland & Yoshikawa, 2013).

Demographics of children enrolled in K1DS classrooms. Table 3 describes the demographic characteristics of the children in the CBO K1DS classrooms in each implementation year. A higher percentage of African American children attended CBO K1DS classrooms than BPS K1. Relatedly, a slightly higher percentage of K1DS students speak English at home than in BPS K1 (67% vs. 50%). Most importantly for the implementation of the OWL and Building Blocks curricula (curricula designed for 4 year olds), not all children in CBO K1DS classrooms were 4-years-old at the beginning of the school year. On average, between 66% (in 2013-2014) and 60% (in 2014-2015) of all children enrolled in a CBO K1DS classrooms were 4 years old. However, it is important to note that classrooms differed widely in the percentage of

children who were 4-years-old, ranging from 12%-100%. The classroom in which 12% of the children were 4-years-old at the beginning of the school year had one 4-year-old child in a classroom otherwise composed of 3-year-old children. One recent study found that four-year-old children scored substantially lower on language, literacy, and mathematics kindergarten readiness assessments when in classrooms with higher percentages of 3-year-olds (Ansari, Purtell, & Gershoff, 2015). Four year old children did not differ in social-emotional skills by classroom age composition.

Only three of the 14 CBO K1DS classrooms enrolled roughly equal proportions of children representing distinct racial/ethnic groups. The remainder of the classrooms enrolled children primarily from one racial/ethnic group. Additionally, over 75% of the children for whom English was not their home language were enrolled in seven of the 14 classrooms. The majority of the children in the remaining seven classrooms were native English speakers.

Table 3 Characteristics of children in CBO K1DS classrooms compared to characteristics of children in the Boston Public Schools K1 program.

Percentage of children who were...	CBO K1DS		BPS K1
	2013-2014	2014-2015	
Latino/a	26%	34%	41%
African American	57%	49%	26%
White	8%	7%	13%
Asian American	6%	6%	11%
Native English Speakers	67%	69%	50%
At least 4-years-old by Sept 1 of school year	66%	60%	100%
Receiving financial assistance to attend prekindergarten (CBO) or receiving free/reduced lunch (BPS)	87%	89%	69%

Note: CBO KIDS 2013-2014: $N_{\text{children}}=259$, CBO K1DS 2014-2015: $N_{\text{children}}=220$. The financial assistance that children in CBO K1DS classrooms received to attend prekindergarten included MA EEC vouchers and UPK subsidies received by the center. BPS K1 $N_{\text{children}}=2,018$. The BPS sample was from the Preparing to Succeed study which included children who attended the BPS K1 Program in 2008-2009 or 2009-2010 (see Weiland & Yoshikawa, 2013). All children in BPS attend for free.

Parent Choice. We surveyed a random sample of CBO K1DS parents in May-June 2015 in order to determine what had led families to select prekindergarten services in the community rather than in the BPS K1 program. Surveys contained a short set of researcher-generated questions asking parents to select the reason or reasons they selected the center in which their

child attended. Parents were also asked whether they had applied for a BPS K1 slot, and if so, whether they had been offered a slot. Parents who had been offered a K1 spot but did not take it were asked a short series of questions regarding their decision not to take the slot. Parents were recruited to take the survey via flyers placed in their children's cubbies or in-person during drop-off. The flyers contained an internet link to a center-specific survey. The response rate was low with only 24% of all possible families responding. Given this, it is difficult to determine whether these data speak to general trends in the CBO K1DS parent group or represent the experiences of a biased subsample.

Overall, 11 of the 48 parent respondents reported applying for a K1 spot in BPS. Of these parents, eight did not receive a spot. Of the three parents who were offered a spot but did not take it, three reported receiving an offer for a different school than desired. These same three parents also reported the operation hours as an issue that factored into their decision. One parent worried about how her child would transition into a new classroom upon receiving a BPS K1 offer after the school year had already begun. Parents reported selecting the CBO center their children attended for a variety of reasons with most parents selecting the cost and hours of operation. Many parents (42% of respondents) also reported that their children had attended the same center since they were infants. See Appendix Table 6 for more details.

The Intervention Components

As the Theory of Change in Figure 2 illustrates, the intervention components consisted of activities organized in five categories. Four of these categories are targeted to the teacher: curricular materials and professional development, coaching, and compensation supplementation. The final category is targeted to center directors: monthly meetings.

Curricula materials

Prior to participating in K1DS, ten classrooms in seven centers used Creative Curriculum, a whole child, child-driven curriculum commonly used in community-based preschool and prekindergarten programs. Two of the remaining centers (two classrooms) were using center-developed curricula. The final two classrooms, located in the same center, were using OWL and Building Blocks prior to participating in K1DS.

With their agreement to participate in K1DS, CBO classrooms were expected to use the literacy and language curriculum OWL and mathematics curriculum Building Blocks. Over the years, BPS has developed specific, day-by-day pacing and planning guides to help K1 teachers integrate both OWL and Building Blocks elements throughout each school day. These were shared with the CBO K1DS teachers, as were formal curriculum guides and other curricular materials from the publishers of each curriculum.

Curricula training

Prior to the beginning of the 2013-2014 school year, CBO K1DS teachers were invited to participate in three large-group OWL and Building Blocks training sessions designed for and attended by newly hired BPS K1 teachers. In the 2013-2014 school year, the intervention model shifted to emphasize mathematics knowledge and instruction. Starting in January 2014, eight 3-hour professional development sessions were offered by district mathematics coordinators for the CBO K1DS teachers. These sessions were called @Scale trainings and focused on building teacher capacity and confidence in mathematics and problem solving. In April 2015, participating classroom teachers and center directors had the option to participate in Touchpoints Training offered by the Brazelton Touchpoints Center at Boston Children's Hospital. This training emphasized building relationships with parents as a way to strengthen support networks for children attending K1DS classrooms.

Coaching

Global support for classroom management, planning, and implementation of the OWL and Building Blocks curricula was provided by a BPS K1 coach from the Department of Early Childhood. As in the BPS K1 curriculum rollout, the coach adapted the content and frequency of her visits to the capacity of CBO K1DS teachers for implementing the two curricula. On average, the coach sought to observe and meet with the teaching team from each CBO K1DS classroom twice per month. The coach was a former preschool teacher who had coached BPS K1 teachers for over five years prior to the start of K1DS.

According to coach interviews, documentation, and teacher interviews, the coaching sessions focused on: behavior management, planning and preparation, integration of Building Blocks activities and pedagogic approaches into all aspects of classroom instructional time, and

understanding and implementing core OWL language and literacy strategies. The coach worked with teaching teams in each classroom to develop an observation and feedback plan that made sense for the level of support teachers needed for different parts of the K1DS model. For example, in some classrooms teachers struggled with managing the disruptive behavior of one or two children, and thus the coach worked with the teaching team on classroom management first and curricula implementation later.

As a companion to the @Scale mathematics professional development sessions, selected K1 teachers from BPS and one CBO K1DS classroom were assigned to mentor CBO K1DS teachers. The mentoring sessions included a classroom observation and a consultation session in which mentors were asked to share best-practices.

Mentors were trained by BPS mathematics education experts to provide targeted and practical advice to teachers regarding their mathematics pedagogy. With the exception of the two mentors from the CBO K1DS center, no mentors attended the professional development sessions. Mentors were assigned to one or two CBO K1DS classrooms and were asked to visit each classroom twice in the spring of 2014. The first visit was to include a classroom observation during which the mentor also modeled high-quality mathematics instructional interactions with children. During the second visit, the mentor was expected to discuss her observations with the classroom teachers and to provide suggestions for improving the integration of mathematics into unstructured moments as well as strengthening mathematics instruction more generally.

Monthly Director's Meetings

Beginning in January 2013, center directors were expected to attend monthly meetings with members of the BPS Department of Early Childhood. These meetings were meant to build community among the participating CBO K1DS leaders and provide directors with an opportunity to ask questions about the K1DS project. In the 2014-2015 school year, directors were also provided with professional development centered on instructional support and measuring quality of mathematics instruction. Toward the end of the program, directors began leading sessions and taking a more active role in agenda setting.

Compensation Supplementation

CBO K1DS teachers received salary and benefits supplementation. Centers were given this supplementation to distribute in a way that made sense within their agency's compensation system. Some teachers received the supplementation as part of their biweekly paychecks while others received one lump sum at the end of each semester. Roughly \$430,000 was distributed to lead and assistant teachers within the participating centers in the period between January 2013 and June 2015. The amount of money teachers received was determined in part by each center's requested amount at the time of application and teachers' education and experience levels. The median amount supplemented was \$10,584 (range of \$3,790-\$14,324 annually), which represented 27% of teachers' median base salary and wages prior to the supplementation. The average K1DS teacher made \$13.83 per hour before K1DS and \$22.89 per hour after the K1DS supplementation. In comparison, Whitebook, Phillips, and Howe (2014) report that nationally, the average hourly wages of prekindergarten teachers across all auspice-type was \$10.33 (range: \$9.20-\$13.90). Massachusetts has long been one of the highest-paying states for ECE teachers; in 2014, preschool teachers in the state on average made \$17.42 per hour (Bureau of Labor Statistics, 2014). We summarize this and provide state and national benchmarks in Table 4.

It is important to note that teachers in the CBO K1DS classrooms were compensated at a much lower rate than BPS K1 teachers. For example, one CBO center advertised the starting salary for a lead teacher at \$31,000-\$35,000 per year and the starting salary for an assistant teacher at \$24,000-\$33,000 depending on the level of educational attainment. In contrast, BPS lists its starting salary for a lead teacher at \$52,600-\$66,500 and the starting salary for an assistant teacher at \$26,300-\$34,900. Even with the highest level of salary supplementation, CBO K1DS teachers earned less than a new K1 teacher in the BPS system regardless of how many years this K1DS teacher had been working in the center and their educational attainment.

Table 4 Average hourly wages of teachers in CBO K1DS, BPS K1, a typical classroom in Massachusetts, and in a typical classroom in the United States.

	Average Hourly Rate	Range
K1DS baseline	\$13.83	\$9.00 - \$20.00
After K1DS supplementation	\$22.89	\$15.00 - \$29.00
BPS K1	\$32.00	\$31.00 - \$33.00
Massachusetts (across auspice type)	\$17.42	--
United States (across auspice type)	\$10.33	\$9.20 - \$13.90

Participation in K1DS Intervention Components

In Table 5, we summarize CBO K1DS director and teacher participation in the program components. On average, directors attended 70% of the monthly project meetings (range of 58-92%). The majority of the original teachers in participating K1DS classrooms attended the large group OWL and Building Blocks training sessions. Teacher attendance at the @Scale sessions was high, averaging 88% (range of 0-100%). The BPS K1 coach met with classroom teachers an average of 16 times in the intervention period. This represents roughly 70% of the total opportunities for monthly coaching. Classroom teachers ranged in their uptake of this intervention component with two classrooms receiving as low as 11 coaching sessions (48%) and one classroom receiving as high as 20 coaching sessions (87%).

When the K1DS program began in January 2013, only teachers from 10 of the 14 classrooms were promised salary supplementation. The salary supplementation for these “round 1” teachers was distributed to centers or their supporting agencies to be distributed in a way that made the most sense given their organizational structures. When interviewed in May-June 2014, teaching teams from 4 classrooms which entered the K1DS program in round 1 were either unaware of or had not received a pay boost. In follow-up investigations in the summer of 2014, we determined that the money had not been distributed in two classrooms (both located in one center) and the amounts requested and paid by the center directors were quite low for the teachers in the other two classrooms. Efforts were made in the 2014-2015 school year to hold centers and their sponsoring agencies accountable for distributing the salary supplementation. All teachers interviewed in May-June 2015 reported receiving additional compensation but were unable to recall the exact amount. Retroactive salary supplementation was received by teachers in the 4 “round 2” classrooms.

Table 5 Teacher and center director participation in key K1DS professional development components during the 2.5 implementation period.

Center	% Director meetings attended	% of OWL & Building Blocks Trainings Attended	% of @Scale Trainings Attended	# Coaching Sessions (January 2013-June 2015)
I	58%	100%	100%	21
II	83%	100%	100%	23
III	50%	100%	100%	18
IV	92%	66%-100%	100%	15
V	75%	100%	100%	20
VI	33%	0% ^φ	0%	22
VII	67%	88%	100%	18
VIII	67%	100%	88-100% [†]	20
IX	83%	100%	100%	20
X	92%	100%	100%	22
Average	70%	88%	92%	20

Note: Ten centers participated in the K1DS program. Four of the participating centers (II, V, VIII, IV) had two classrooms participating in the pilot. The other 6 centers had one classroom each. The % of teacher trainings attended column indicates what percentage of all offered sessions were attended by at least one teacher from each classroom participating in K1DS within each center. ^φ=This center entered the K1DS program in the second round and were not integrated into the OWL and Building Blocks training session. While already a part of K1DS for an entire year, no teachers from this center attended the @Scale trainings. [†]= There were two classrooms in this CBO K1DS classroom. The teachers in one classroom missed one of the eight @Scale sessions (attending 88%) whereas the teachers in the other classroom were present for all of the @Scale trainings.

Methods for Collecting Data and Evaluating Progress

Child sample recruitment. Parental consent forms were available in Spanish and English. Consent forms for all adult participants (including parents, teachers, directors, and BPS staff) were only available in English. All teachers and directors completed consent forms. All families who had a child enrolled in a CBO K1DS classroom at the beginning of each full school year (October 2013 or October 2014) were asked to sign a consent form.

CBO K1DS center directors and teachers worked with evaluation staff to recruit children into the study. Roughly 80% of all children initially enrolled in each participating classroom were consented each school year. As many classrooms served 3-year-olds as well as 4-year-olds, we consented both children who were too young to have entered BPS K1 and children who would have qualified for BPS K1. One challenge we faced in gathering data on the children enrolled in the K1DS classrooms was a trend in many centers of unstable enrollment. Five of 14 classrooms reported low initial enrollment that fluctuated through the first four months of each

academic year. That is, in these five classrooms, the set of children who “belonged” in the classroom did not stabilize until December or January. In interviews with teachers from these classrooms and center directors, the instability is due in part to 1) the movement of parents out of the community programs when BPS K1 spots become available in the early months of the school year, 2) mid-year expiration of child care vouchers provided by the state of Massachusetts, and 3) low numbers of 4-year-olds who need placement in a CBO prekindergarten classroom relative to past years. Directors attribute (3), in part, to the growth of the BPS K1 system from 2005 to 2013. In the appendix, Table 4 summarizes demographic characteristics of our full sample of 4-year-olds compared to the characteristics of the children who were assessed in both fall and spring of the school year. Children were similar on all characteristics except receipt of financial assistance to attend prekindergarten (70% of those who participated received financial assistance vs. 87% in the full sample of CBO K1DS children). We view this as an indication that children whose families received state-assistance may have dropped out from the fall to spring assessment period.

In cohort 1 (2013-2014), 130 of the 197 consented children were 4-years-old and 120 of these children were assessed in the fall of 2013. One hundred children (or 73% of the 130 consented children in cohort 1) were assessed in both the fall of 2013 and the spring of 2014. In cohort 2 (2014-2015), 120 of the 200 consented children were 4-years-old and 110 of these children were assessed in the fall of 2014. Sixty-three children (or 53% of the original 120 4-year-old children) were assessed in both the fall of 2014 and the spring of 2015. One difference between the two cohorts was the timing of the consent and fall assessment period. In cohort 1, we began the consent and assessment process concurrently in late September and early October whereas in cohort 2 parents were asked to sign consent forms upon enrollment in early summer through August and assessments occurred in early October.

Overview of methods used and data collection timing. We employed a range of methods to gather data from all K1DS program stakeholders. Table 6 summarizes the data collection methods and timing for the evaluation of K1DS. In the rest of this section we discuss the instruments used in the evaluation of the CBO K1DS project.

Table 6 Timing of measurements by stakeholder and data collection type for the K1DS study.

	January 2013 (Baseline)	Fall 2013	Spring 2014	Fall 2014	Spring 2015
Surveys		Directors, Teachers			Parents
Interviews		BPS Department of Early Childhood staff; K1DS funding partners; BPS K1 coach; Directors; Teachers;	Coach, Directors, Teachers	Coach	Coach, Directors, Teachers
Classroom observations	CBO K1DS (N=13) & BPS K1 (N =9)		CBO K1DS (N =14)		CBO K1DS (N=10) & BPS K1 (N =23)
Child Assessments		CBO K1DS	CBO K1DS	CBO K1DS	CBO K1DS
Fidelity Observations					CBO K1DS
Document Review		❖	❖	❖	❖

Key ❖= documents reviewed included: original RFP, notes from coaching logs, publicly available documentation about CBOs; and press coverage of project and CBOs. “Coach” refers to the BPS K1 coach working with CBO K1DS teachers.

Measures

Surveys. Surveys were administered to teachers and directors in September-October 2013 (e.g., the start of the first full school year of the intervention). The surveys were administered by the evaluation director following the fall 2013 interview. Teachers and directors had the opportunity to complete the survey online or on paper. These surveys sought to capture respondents’ professional training and experience. They also sought to quantify participation rates in and attitudes toward K1DS professional development activities that had occurred from January 2013-September 2013.

These surveys also determined respondents’ perceptions of their center’s readiness for the K1DS intervention. This was done by including an Assessment of a (pre)School’s Readiness for Change (Wanless, 2013). This survey targets 4-levels of characteristics that Wanless (2013) hypothesized contribute to a center’s capacity and willingness to engage with an intervention or other kind of new programming. The 4-levels are: community, center, director, and teacher.

Respondents are asked to what degree they agree (1=strongly disagree, 3=neither agree nor disagree, 5=strongly agree) with 24 statements. The community subscale consists of 4 items ($\alpha = 0.54$) including: *“This (pre)school has worked with at least a few other local organizations (related non-profit organizations, other schools, universities, etc.) in the past”* and *“In this community, there are currently new initiatives, funding opportunities, and/or professional development experiences being offered on the topic of this intervention (SEL, math, etc.)”*. The director subscale also consists of 4 items ($\alpha = 0.70$) which include: *“Teachers in this (pre)school generally believe that their principal/director genuinely believes in the critical importance of the topic of the intervention (SEL, math, etc.)”* and *“In general, teachers in this (pre)school feel like their director/principal is “on their side.”* The teacher subscale consists of eight items ($\alpha = 0.78$) including: *“People who have observed me teaching would say that I enjoy teaching children”* and *“I seek out information about new teaching strategies that might benefit my children.”* And finally, the center subscale consists of eight items ($\alpha = 0.79$) including: *“Teachers in this (pre)school feel like the amount of stress in their job is manageable”* and *“This (pre)school’s schedule allows time for teachers to exchange ideas.”* All directors completed the surveys and 86% of CBO teachers (representing 13 out of 14 classrooms) completed the survey.

In May-June 2014, questions that had previously been part of the director and teacher survey were integrated into the interview protocol to reduce the burden on study participants.

Interviews. As Table 6 illustrates, we interviewed stakeholders from all organizations involved with the CBO K1DS project. We interviewed the Director of the Department of Early Childhood at BPS, the BPS School Readiness Manager, and the BPS-based K1DS project director once in the fall of 2013. These interviews solicited information about the professional training and experience of the respondents as well as information about the motivation and history of the BPS K1 program and the CBO K1DS project. In the fall of 2013 we also interviewed the leadership team of Boston Thrive in 5 ($N=3$). These interviews focused on the motivation and history of the BPS K1 program, the CBO K1DS project, and parallel efforts in Boston to improve quality of education and care services to Boston’s young children.

CBO K1DS center directors and teachers were interviewed three times in the study period. The first interviews were conducted in September-October 2013. All K1DS teachers

($N=34$) and directors ($N=10$) plus three CBO center-based instructional leaders were interviewed. These interviews complemented the surveys used in this time frame and sought respondents' motivation for working with young children, hopes and worries about the K1DS project, perceptions of the strengths and challenges of their organization, and perceptions of the K1DS intervention components. The second round of interviews occurred in May-June 2014 and asked follow-up questions derived from data collected in fall 2013. During this round of interviews, two teachers from each classroom ($N=28$), all center directors ($N=10$), and two instructional coaches were interviewed. These interviews also solicited respondents' perception of how the K1DS project had worked for them in the first full school year of implementation with questions about each of the intervention activities included in the interviews. The third round of interviews occurred during May-June 2015. In this round, two teachers in 14 classrooms were interviewed. In the classrooms in which turnover had occurred or coverage issues prevented teachers from leaving the classroom during the window for interviews ($N=4$ classrooms), 1 teacher was interviewed per classroom. Most center directors ($N=8$) were interviewed about their work with the K1DS project in the final year of implementation of the project. These interviews focused on respondents' experience of implementation in the second full school year and asked participants to reflect upon the similarities and differences in their experiences from school year 1 to school year 2. Respondents were also asked to reflect on the challenges they faced in implementation, share examples of success, and describe their perceptions of any differences in child learning that had occurred as they integrated the OWL and Building Blocks curricula into their classroom practice.

The K1DS coach was interviewed six times throughout the course of the 2.5-year project—September 2013, October 2013, April 2014, July 2014, December 2014, May 2015. During these interviews, the coach was asked to describe her process of working with the CBO K1DS directors and teachers and well as share her notes and perceptions of the implementation processes occurring in the classrooms. During each interview, the coach was also asked to reflect on the successes and struggles of each classroom and center.

CBO Classroom Observations. CBO K1DS classrooms were observed four times. Three of these times (January 2013, May-June 2014, and May-June 2015) focused on observing and

rating language, literacy, and mathematics instructional quality. Each observation began after children in the CBO K1DS classrooms had completed breakfast and ended when they started lunch, lasting approximately 3-4 hours. The fourth observation (April-May 2015) focused on observing and measuring the fidelity of implementation of the OWL and Building Blocks curricula in the classrooms. The timing and length of these observations was similar to that of the classroom quality observations. Teachers and center directors were notified in advance of the purpose and date of the observations. In all cases, observers who were trained to be reliable on the instruments conducted the observations with validated observation protocols (80% minimum agreement standard, with ratings counted as “agree” if they were within 1 point). Below, we describe the four observation protocols used in this study.

The first observation protocol was the Early Language and Literacy Classroom Observation tool (ELLCO; Smith et al., 2002). The ELLCO was used in all three classroom quality observations (January 2013, May-June 2014, and May-June 2015). The ELLCO is used to evaluate the environment and teaching practices of classrooms in regards to language and literacy. It is comprised of three scales: a literacy environment checklist, a general classroom environment quality scale, and a language, literacy, and curriculum quality scale. We used the eight-item literacy activities rating scale, as in Burchinal, Xue, Tien, Auger, and Mashburn (2011). We also used the general classroom environment scale. Ratings on these scales are determined as follows: a score of five is considered “exemplary,” three is “basic,” and one is “deficient.”

The second observation protocol, the Classroom Observation of Early Mathematics—Environment and Teaching (COEMET; Samsara & Clements, 2009)—evaluated the number, nature, and quality of mathematics instructional moments in prekindergarten classrooms. The COEMET was used in all three classroom quality observations (January 2013, May-June 2014, and May-June 2015). We used two scales from the COEMET. The first one, classroom culture, is a measure of how much mathematics imagery and language permeates the physical environment of the classroom. The second scale is a measure of the quality of in-depth mathematics instruction. Instances of in-depth mathematics instruction are called *specific math activities or SMAs* in the COEMET manual. The average quality of SMAs refers to the quality of

small-group mathematics activities and instruction. These are the more formal, more complex, and intentional aspects of mathematics instruction occurring in the classroom (versus simple math songs or calendar activities). We also counted the number and timed the average length (in minutes) of in-depth math activities.

The third observation protocol was the Classroom Assessment Scoring System (CLASS; Pianta, La Paro, & Hamre, 2008). The CLASS was used in the last classroom quality observation (May-June 2015). The CLASS focuses on interactions between students and teachers. It is measured on a seven-point scale with higher scores indicating higher quality. A score of seven is considered “excellent,” five is “good,” three is “minimal,” and one is “inadequate.” The CLASS consists of three subscales – Emotional Climate, Instructional Support, and Classroom Organization – and each subscale is composed of multiple indicators. In our study, observations for the CLASS were conducted in 30-min cycles for observation and scoring.

The fourth observation protocol used was comprised of two tools: the Building Blocks Near Fidelity tool (Sarama, et al., 2012) and a streamlined version of the OWL fidelity tool used in a previous BPS study (Weiland & Yoshikawa, 2013). The Building Blocks measure included a general curriculum section, along with sections that focused on the implementation of specific components of the curriculum. The OWL measure included separate sections and items for each component of the OWL curriculum. In both tools, items were scored either dichotomously (yes/no) or using a five-point Likert scale (where 1=almost never, 2=rarely, 3=sometimes, 4=usually, 5=almost always) to record whether that particular component of the curriculum had been implemented in the classroom. We constructed summary measures of adherence to curricula by curriculum component (e.g., separate scores for Intro to Centers, Centers, Storytime, in OWL). We added up the ratings across items for each component and then divided by the maximum score for that component. As an example, if a component had 10 items, all with a 0-4 rating scale, a classroom’s adherence score could have fallen between 0 and 100 (0/40 to 40/40). Adherence scores accordingly can be viewed as the percentage of full implementation achieved (e.g., 50% level of implementation for a given component). For each curriculum, we created a summary score by averaging adherence scores for all observed components for that curriculum.

Implementation fidelity was observed separately from classroom quality. Classrooms were co-observed on fidelity by the BPS K1DS coach and an external observer.⁵ Classrooms were observed on a typical day in the morning, with a 9:30 AM start. Observers stayed through lunch at each site, which typically meant they observed until noon or 12:30. They rated any OWL and Building Blocks components that were observed within this window. After each observation, the two observers compared ratings, discussed any disagreements, and in those cases agreed on a final code.

Time use in the classroom and time in whole-group activities was also tracked during the fidelity observation. During each fidelity observation, the external observer kept track of the dominant activity/configuration in each classroom. An activity was considered to have ended and another begun depending on the teachers' behavior (e.g., once the teacher finished letting children choose centers and transitioned to center time, center time was considered to have begun and morning meeting time to have ended).

BPS Classroom Observations. For baseline in January 2013, we used a BPS tool (discoverbps.org) to identify the two BPS schools with K1 programs closest to each CBO center. For each center, we randomly selected a K1 classroom geographic control. Three randomly selected teachers did not agree to participate and were replaced with other randomly selected teachers for that center-BPS match. Observers then conducted ELLCO and COEMET observations per the same method in CBO classrooms.

In spring 2015, 23 K1 classrooms were observed as part of the department's biannual quality monitoring process. The study team recruited the sample in two stages: first schools, and then classrooms within schools. The school-level agreement rate was 87.5% and for K1 teachers in those schools, 78.8%. Within participating classrooms, the same spring 2015 K1DS observers conducted the same battery of assessments as they did in the CBO K1DS classrooms: CLASS, COEMET, and ELLCO.

Child Assessments. All children for whom we obtained parental consent were assessed at the beginning and end of each academic year in the K1DS project (two cohorts total). Six

⁵ Due to scheduling difficulties, the BPS K1DS coach observed one classroom on her own.

instruments were used by assessors trained to be reliable during 12 hours of small group training and practice and one on-site reliability check. The assessors were all former early childhood teachers. Children were assessed individually with assessment sessions lasting between 20-45 minutes each. In the rest of this section, we describe the six instruments used to assess children's learning and development. The assessments were administered to individual children by trained child assessors.

Children's receptive vocabulary was measured using the Peabody Picture Vocabulary Test III (PPVT-III; Dunn & Dunn, 1997), a nationally normed measure that has been used widely in diverse samples of young children (U.S. Department of Health and Human Services, 2010). The test has excellent split-half and test-retest reliability estimates, as well as strong qualitative and quantitative validity properties (Dunn & Dunn, 1997). It requires children to choose (verbally or nonverbally) which of four pictures best represents a stimulus word. In our analysis, we used the raw score total in our study.

The second instrument was the Woodcock-Johnson Letter-Word Identification subscale (Woodcock, McGrew, & Mather, 2001). This instrument is a nationally normed, widely used measure (Gormley et al., 2005; Peisner-Feinberg et al., 2001) of children's early literacy skills. Children are asked to identify and pronounce isolated letters and entire words fluently. According to the developers, the estimated test-retest reliability of the Letter-Word subscale for 2- to 7-year-olds is 0.96. Consistent with other prekindergarten studies (Gormley et al., 2005; Gormley et al., 2008), we used the raw score total.

The third instrument was the Woodcock-Johnson Applied Problems subscale (Woodcock et al., 2001) which assesses the numeracy and early math skills of children by requiring them to perform relatively simple calculations to analyze and solve arithmetic problems. Its estimated test-retest reliability for 2- to 7-year-old children is 0.90 (Woodcock et al., 2001) and it has been used widely with diverse populations of young children (Gormley et al., 2005; Peisner-Feinberg et al., 2001; Wong et al., 2008). We used the raw score total.

The fourth instrument measured a principal dimension of executive functioning called working memory. This instrument, the Forward Digit Span (FDS; Gathercole & Pickering, 2000; Wechsler, 1986) measures phonological loop. In this task, the assessor reads aloud a string of

numbers to the test child, with approximately a 1-second pause between digits. The child then has to repeat back exactly what the assessor said. Before items are administered, the child must pass a practice trial, demonstrating that he or she understands the directions of the task. FDS is scored from 1 to 6. The score represents the child's digit span memory (i.e., a 2 represents a digit span memory of two digits).

The fifth instrument tapped into a principal dimension of executive functioning called cognitive inhibitory control. This instrument is called Pencil Tapping (Diamond & Taylor, 1996). During this assessment, the child was asked to tap twice if the evaluator tapped once and tap once if the evaluator tapped twice. Assessors first administered a set of practice trials to ensure that children understood the rules of the task. Children who passed the practice trials were then administered 16 total trials. The task measures children's cognitive inhibitory control and, to a lesser degree, working memory and fine motor activity (Bierman, et al., 2008). Scores recorded the correct number of trials out of 16 that children achieved. Because of concern that tapping a pencil could prove difficult for preschoolers and might conflate cognitive inhibitory control with fine motor skills, we substituted larger wooden kitchen spoons for pencils in this task.

We used the sixth and final instrument, called the Task Orientation Questionnaire (TOQ; Smith-Donald, Raver, Hayes, & Richardson, 2007) to measure three different skills. The full TOQ assesses the child's emotional state and capacity to sustain focus on a set of tasks during a testing session. After administering the child assessment battery, assessors rated each child on 13 items reflecting his or her capacity to sustain attention to the tasks, demonstrate self-regulation, and engage actively to achieve a goal. Each item was rated on a 4-point scale, with clear behavioral descriptors provided for each point on the scale. We created three composites to measure our central constructs using the strategies developed in the Preparing to Succeed study (Weiland & Yoshikawa, 2013). The attention shifting composite consisted of four items: "*Pays attention to instructions and demonstration,*" "*Careful, interested in accuracy,*" "*Sustains concentration—willing to try repetitive tasks,*" and "*Cooperates, complies with tester's requests.*" In our analyses, we used a unit-weighted average of responses to these four items.

In the Preparing to Succeed study, Weiland and Yoshikawa (2013) conducted a confirmatory factor analysis with data from the TOQ and identified three items for positive emotion: *“alert and interactive; is not withdrawn,” “shows pleasure in accomplishment and active task mastery,”* and *“confident”*; and three items for impulse control: *“can wait during and between tasks,” “remains in seat appropriately during test,”* and *“modulates and regulates arousal level in self.”* In our analyses, scores on our Positive Emotion and Impulse Control outcomes were unit-weighted averages of children’s responses to the position emotion.

Document Review. To round out our understanding of important aspects of the K1DS program, we collected lesson plans, field notes from our own and the coach’s visits to classrooms and centers, original application materials, field notes from directors’ meetings, publicly available documentation about the agencies and organizations involved in the K1DS project, the original RFP to programs, and all newspaper and marketing material generated by the district or partners throughout the course of the pilot.

Analytic Approach

Given the non-experimental design of this evaluation, we used descriptive strategies to investigate research question 1 (Did implementing the Boston model in CBO K1DS classrooms improve instructional quality?), 2 (How did the mathematics, language, and literacy instruction in CBO K1DS classrooms change?), and 4 (Did children enrolled in CBO K1DS classrooms show gains in their school readiness skills?). Specifically, we examined change over time in the means of quantitative assessments collected at each time point for CBO K1DS classrooms and child assessment scores; compared CBO means to means from BPS K1 samples collected for the present study sample and in past data collection efforts in BPS K1 (Weiland et al., 2013; Weiland & Yoshikawa, 2013); and, where appropriate, conducted t-tests of these observed mean differences. We also estimated bivariate correlations to determine the magnitude and statistical significance of relationships between: 1) fidelity ratings and classroom and center characteristics; and 2) classroom quality and classroom and center characteristics.

In order to conduct our descriptive analyses and address research question 3 (What were the barriers to implementing the Boston model in CBO K1DS classrooms?), we developed and employed a range of strategies for working with qualitative data. As interviews were

conducted, researchers wrote analytic memos capturing the essential points of the interview and noting any follow-up questions. From these memos, a preliminary coding system was generated for each of the following themes: challenges, supports, experiences. Interviews, which had been audio recorded, were transcribed and entered into a qualitative analysis software package called NVivo (NVivo qualitative data analysis Software; QSR International Pty Ltd. Version 10, 2012). The codes generated in the analytic memo writing stage were used to create course categories of challenges, supports, and experiences. In this way, we were able to capture the range of unique entries for each category and count instances in which the same challenge, support, or experience was mentioned across more than one respondent or within respondents over time (Maxwell, 2012; Lin, 2009).

Findings

Teacher and Director Readiness to Change

CBO K1DS center director and teacher perceptions of their center's readiness for the changes entailed in the K1DS project were solicited during the fall 2013 survey, using the Readiness to Change scale (Wanless, 2013). Teacher and director perceptions of readiness to change were measured overall and in regard to four different levels. Each composite is scaled from 1-5. We show the average ratings of teachers and directors on each of the subscales of readiness to change in Table 7. Average ratings are quite high (with a value of 4 or higher on the 5-point scale) with the only low rating being that of directors regarding their own teachers' readiness to change. Teachers rated their own readiness to change highly.

Table 7 Average ratings on four subscales and an overall score from the Readiness to Change scale for teachers, directors, and for the full sample. Bivariate correlations between teacher and director ratings on each subscale as well as p-values.

	Teacher Mean (SD)	Director Mean (SD)	Full sample Mean (SD)	Correlation	p-value
Community	4.06 (0.40)	4.53 (0.37)	4.25 (0.45)	0.72	* $p = 0.04$
Director	3.88 (0.71)	4.23 (0.51)	4.03 (0.65)	0.51	$p = 0.20$
Teacher	4.15 (0.48)	4.61 (0.24)	4.34 (0.45)	0.20	$p = 0.63$
Preschool	4.23 (0.47)	4.59 (0.19)	4.38 (0.41)	0.40	$p = 0.33$
Overall	4.12 (0.47)	4.52 (0.20)	4.29 (0.43)	0.65	$p = 0.08$

Notes: 29 teachers completed the Readiness to Change survey items, with representatives from all 10 centers. Teacher ratings were averaged within centers to create center-level teacher ratings of readiness for change that were then compared to director ratings. All ten center directors completed the Readiness to Change survey items.

Intervention components were well-received by K1DS participants

Generally, directors were pleased to have the opportunity to build a professional community with directors from across the Boston area. Directors and teachers universally agreed that the coaching portion of the K1DS program was a major benefit of the program. While some participants requested specific kinds of changes to the coaching model, including more information sharing between the coach and the center director (4 out of 10 directors requested this during the May-June 2014 interview), all described specific ways in which the coaching professionalized the work of teachers and enhanced planning for and implementation of the OWL and Building Blocks curricula.

Feedback on the mathematics professional development sessions was mixed. Surveys administered directly after the training sessions yielded positive responses to prompts such as *“I will use what I learned in this professional development session this school year”* and *“I learned enough to make this professional development session worth my time.”* The average response to these questions across all eight sessions was equal to 4.5 or above on a 5-point scale (5=strongly agree). However, when interviewed, 80% of teachers reported being frustrated by the pace and content of the sessions.

The trainings were planned were planned to increase teacher mathematics knowledge, in accordance with a theory that increased teacher mathematics knowledge is necessary for improving math instruction. However, teachers were less interested in increasing their own knowledge of and comfort with mathematics than they were in gaining concrete and directly

applicable knowledge about teaching mathematics to young children. This disconnect was at the heart of many teachers' disappointment with these sessions.

Fidelity to curricula

Dosage, as measured by the fidelity instrument. CBO K1DS classrooms have a longer day than BPS, with some running from 8:00 am – 6:00 pm compared to the BPS 6-6.5-hour school day. This is a potential strength of CBOs, as more total time presents the opportunity for greater curricula dosage.

Using data from observer notes during fidelity observations in CBO classrooms in spring 2015, we found that approximately 53% of the time, or 80 out of 150 minutes, observed was spent on instruction (range=43-73%) and 37% of instructional time was spent in whole group instruction (range=0-60%). Classrooms spent between 9% to 32% of the time on transitions, for an average of 17% of time on transitions.

The overall pattern across K1DS classrooms was that not enough time was devoted to OWL and Building Blocks. In total, core OWL curriculum components require approximately 3 hours per day to implement fully. For Building Blocks, when fully implemented, children are exposed to approximately 180 minutes of Building Blocks per week, through whole group, small group, center and computer activities.⁶ From K1DS classroom schedules, most teachers delivered the core of their Building Blocks and OWL instruction in the morning. Some classrooms did implement some OWL/Building Blocks components in the afternoon – generally a half-hour of additional curriculum time. Adding this 30-minute block to the 80-minute average, CBO classrooms on average failed to reach the benchmarks for curriculum dosage.

In interviews with teachers, we learned that children's staggered drop-offs and pick-ups was a barrier in devoting adequate time to curriculum delivery. District K1 classrooms follow a traditional school-day schedule, with children in school for the same consistent schedule (e.g., 8:30 am start for all students and 3:00 pm end). Though the CBO day was as long as 8:00 am –

⁶ In BPS's implementation of the two curricula together, Building Blocks is often delivered in conjunction with OWL structures – e.g., Building Blocks and OWL small groups run simultaneously (teacher leads one, while the assistant leads the other). According, expected total time on OWL and on Building Blocks in BPS is not the same as expected total time when each curriculum is implemented without the other. The time benchmarks presented here are estimates.

6:00 pm in some centers, CBO teachers generally waited until 9:30 to begin core instruction. It was at this point that breakfast was completed and when most of the children would have arrived. Pick-ups were staggered throughout the afternoon and varied for families from day-to-day. This inconsistency limited the amount of instruction that could be delivered in the afternoon.

Math dosage as measured by the quality observations. In observations of math quality at baseline, spring 2014, and spring 2015, observers recorded several indicators of the quantity of math instruction. Table 8 summarizes this information and compares CBO K1DS and BPS classrooms on these metrics. On average at baseline, CBO classrooms implemented 2.4 in-depth math activities, lasting 6.5 minutes each, and 2.9 routine math activities (e.g., songs, counting number of kids present, etc.). CBO K1DS teachers slightly reduced the number of activities implemented over time. Average time on in-depth math activities increased from baseline to spring 2014 but dipped slightly below baseline levels in spring 2015.

Comparisons to BPS K1 classrooms in Spring 2015 revealed CBO K1DS classrooms spent less time on in-depth math activities at all time points than BPS K1 classrooms; implemented fewer in-depth math activities at all time points; and implemented more routine math activities. Multiplying the average number in-depth math activities by the average length of these activities reveals 16 minutes at baseline, 20 minutes in Spring 2014, and 12 minutes in Spring 2015 spent on in-depth math instruction total in CBO K1DS classrooms compared to 29 minutes in BPS K1 in the spring of 2015. The last three columns of the table present standardized differences between CBO and BPS classrooms. These metrics reveal that some differences were small (e.g., in-depth math activity length in spring 2014) and others were large (e.g., in-depth math activity length in spring 2015).

Table 8 Math instructional quantity in CBO K1DS classrooms over time and in comparison to BPS K1 classrooms.

	CBO			BPS Spring 2015 (N=23)	CBO-BPS diffs. (standardized)		
	Baseline (N=13)	Spring 2014 (N=14)	Spring 2015 (N=10)		Baseline	Spring 2014	Spring 2015
N in-depth math activities	2.40 (1.70)	2.29 (2.64)	1.90 (1.20)	3.22 (2.04)	-0.40	-0.46	-0.65
Average length (in min) of in-depth math activities	6.51 (3.13)	8.75 (7.95)	6.20 (4.42)	9.01 (4.78)	-0.52	-0.05	-1.38
N routine math activities	2.90 (1.7)	2.29 (1.77)	2.60 (1.71)	2.57 (2.41)	0.14	-0.12	0.01

Note: CBO K1DS-BPS K1 differences were standardized by dividing their difference by the relevant BPS K1 Spring 2015 standard deviation. CBO K1DS means were not statistically significantly different, nor were CBO K1DS-BPS K1 differences in means ($p>0.05$).

Adherence. We also observed classrooms for adherence to curricula components (e.g., when implementing centers, were they implementing using OWL teaching strategies, structures, and materials?). Table 9 summarizes these results, by curriculum and component and across all components within curriculum.

Column three of the Table 9 reports in how many centers a given component was observed. On average, combined across the two curricula, observers saw 4.8 components (out of a possible 12) per classroom during the fidelity observations (range=2 to 7). For OWL, the socio-emotional component Let's Talk About It was the least commonly observed component (0/14 classrooms) and Intro to Centers the most commonly observed (13/14 classrooms). For Building Blocks, centers was least commonly observed (1/14) and small groups was most commonly observed (10/14).

In column 4, we report average adherence scores, scored from 0% (no adherence) to 100% (perfect adherence). For OWL, overall adherence was 56% across all components and 59% for just those components most commonly observed across classrooms (Intro to Centers, Centers, and Storytime). In the classrooms in which a given component was observed, Small Groups was the best implemented component (75%, observed in $N=4$ classrooms). Intro to Centers was the least well implemented component (39%, across 13 classrooms). For Building Blocks, overall adherence was 51% across all components and 59% for just those components most commonly observed across classrooms (general, whole group, small group). In the

classrooms in which a given component was observed, whole group was the best implemented Building Blocks component (62%, observed in $N=8$ classrooms). Computers was the least well implemented component (27%, observed in $N=2$ classrooms).

Table 9 Adherence to OWL and Building Blocks Curricula in 14 CBO classrooms

Curriculum	Components	N (%) in which component was observed	Avg. score	SD	Min	Max	N meeting benchmark of 65%
OWL	Intro to Centers	13 (92.9)	38.5	23.7	0	72.2	2
	Centers	12 (85.7)	55.0	16.4	39.6	79.2	4
	SWPL	4 (28.6)	65.6	5.1	59.4	71.9	3
	Small Groups	4 (28.6)	75.0	30.0	30	90.0	3
	Storytime	11 (78.6)	51.7	12.5	38.5	79.2	2
	Let's Find Out About It	2 (14.3)	70.8	1.2	70	71.7	2
	Let's Talk About It	0 (0)	--	--	--	--	--
	<i>Across all components</i>	<i>14 (100)</i>	<i>51.3</i>	<i>17.1</i>	<i>22.1</i>	<i>77.0</i>	<i>4</i>
<i>Across most commonly obs components (Intro, Centers, Storytime)</i>	<i>9 (64.3)</i>	<i>46.7</i>	<i>15.0</i>	<i>36.3</i>	<i>73.2</i>	<i>1</i>	
Building Blocks	General	12 (85.7)	55.9	27.4	15	90.0	6
	Centers	1 (7.1)	50.0	--	--	--	0
	Whole group	8 (57.1)	61.7	18.6	28.6	85.7	5
	Small groups	10 (71.4)	54.0	16.7	36.9	86.9	2
	Computers	2 (14.3)	26.9	27.2	7.7	46.2	0
	<i>Across components</i>	<i>12 (85.7)</i>	<i>55.9</i>	<i>18.4</i>	<i>21.9</i>	<i>82.3</i>	<i>4</i>
	<i>Across most commonly obs components (General, Whole Group, Small Groups)</i>	<i>7 (50.0)</i>	<i>59.1</i>	<i>18.5</i>	<i>34.1</i>	<i>82.7</i>	<i>3</i>

Note: Fidelity components were scaled from 0-100, where 0 represents the lowest score possible on all items for that component and 100 represents the highest score possible on all items for that component. A score of 50 represents a 50% level of implementation. Combined across the two curricula, on average, observers saw 4.8 components per classroom during the fidelity observations (range=2 to 7).

The final column in Table 9 details the number of classrooms (out of 14) that met the benchmark of 65% adherence for that component. Across components and curricula, a small number of classrooms met these benchmarks (e.g., low of 0 for Building Blocks centers and computers and high of 6 for Building Blocks general scale). Table 10 presents another metric for examining adherence in terms of benchmarks, with cutoffs for high, medium, and low

adherence for each curriculum and across the two. Here, three classrooms were high implementers, seven were in the medium range, and four were low implementers.

Table 10 Classroom-level fidelity to curricula in CBO classrooms, expressed in terms of meeting benchmarks

School	Teacher	OWL benchmark met	Building Blocks Benchmark met	Overall Fidelity benchmark met
1	1	High	High	High
2	2	High	High	High
2	3	High	High	High
3	4	Medium	High	Medium
4	5	Medium	Medium	Medium
5	6	Medium	Medium	Medium
6	7	Medium	Medium	Medium
7	8	Medium	Medium	Medium
8	9	Low	Medium	Medium
4	10	Medium	--	Medium
5	11	Low	Low	Low
9	12	Low	Low	Low
9	13	Low	Low	Low
10	14	Low	Low	Low

Note: 65%+ is high fidelity; >40% is medium fidelity; <40% is low fidelity. Cutoffs were empirically determined for the purposes of our study (e.g., they are not prescribed by the OWL or Building Blocks developers).

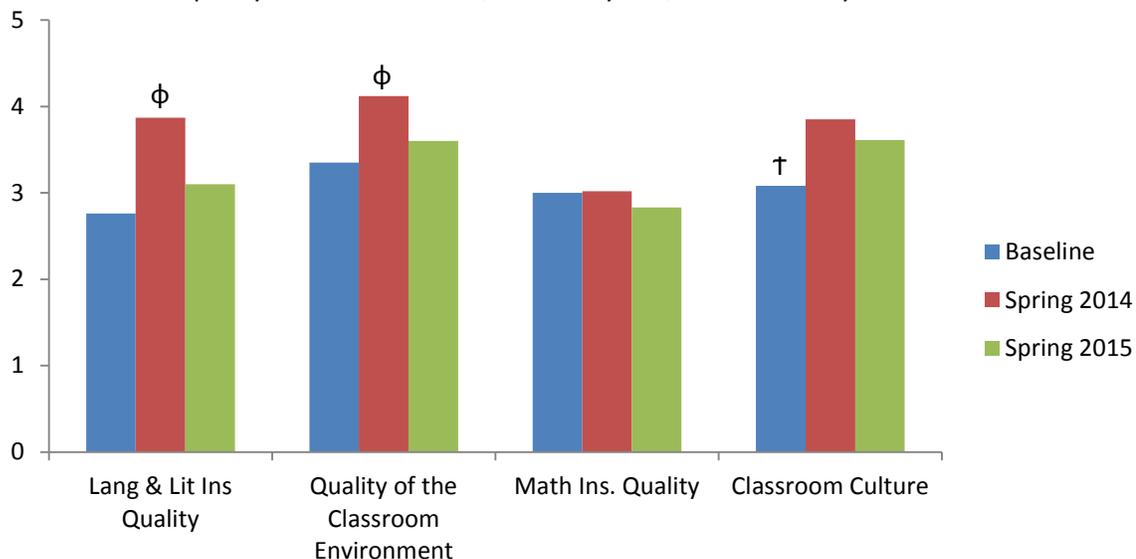
Classroom Quality

Classroom quality was measured at three time points in the K1DS study (January 2013, Spring 2014, and Spring 2015). This section describes the changes in classroom quality observed over these three measurement instances as well as compared to CBO K1DS classroom quality to the classroom quality observed in BPS K1 classrooms in Spring 2015. For comparisons of CBO K1DS classroom quality to other BPS K1 classrooms observed at other time periods (Spring 2010 and Winter 2013), please see the Appendix Table 2.

Figure 5 shows CBO K1DS classroom math, language, and literacy instructional quality across the three time points. At baseline, CBO K1DS classrooms on average scored near the “adequate benchmark” across these metrics. In spring 2014, scores improved to roughly the “good” benchmark for three out of four scales and stayed about the same for math

instructional quality. In spring 2015, scores had declined for all four metrics compared to spring 2014, though they were higher than the baseline scores for 3 out of 4 scales.

Figure 5 CBO classroom quality scores at baseline, after 1.5 years, and after 2.5 years

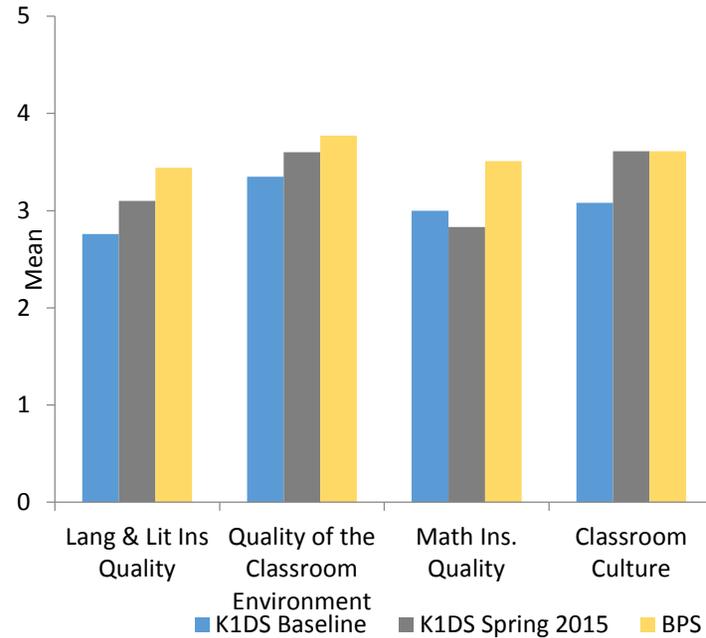


Note: $N=13$ at baseline; $N=14$ in Spring 2014; $N=10$ in Spring 2015. Language and Literacy Instructional Quality and Quality of the Classroom Environment are subscales of the ELLCO. Math Instructional Quality and Classroom Culture are COEMET subscales. ϕ : Spring 2014 scores for both ELLCO subscales were statistically significantly different from scores at baseline and in Spring 2015 ($p<.05$). τ : Baseline and Spring 2014 scores on Classroom Culture were statistically significantly different ($p<.05$). Other within-scale means shown in the figure were not statistically significantly different.

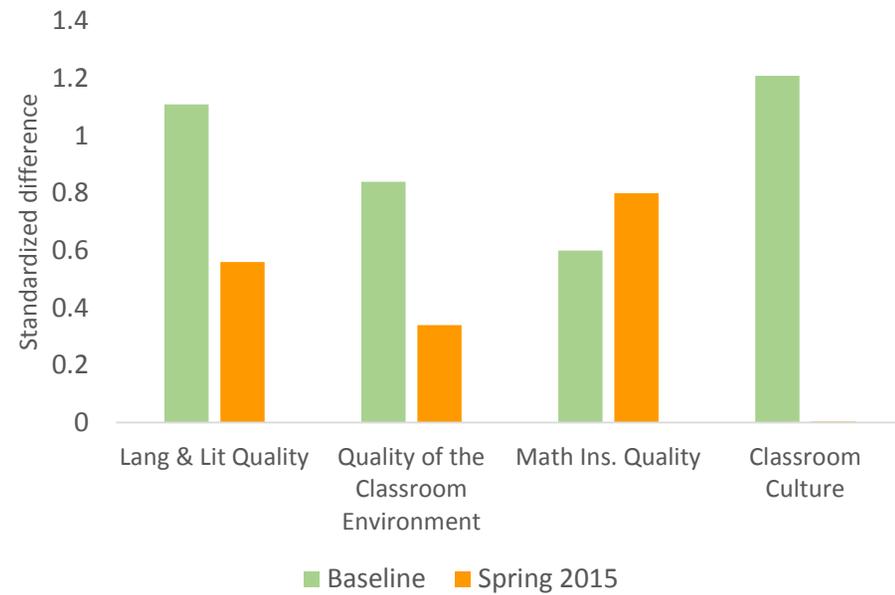
At the start of the K1DS project, CBO K1DS classrooms lagged behind BPS K1 classrooms in language and literacy and mathematics instructional quality (see Figure 6 Panel A). All of the differences were statistically significant. The standard deviations of these measures is generally small (1 point or less). Accordingly, in Figure 6 Panel B, we translated them into standardized difference (or standard deviation) units; the height of each bar indicates how much lower the quality was in K1DS classrooms compared to the average quality in K1 classrooms before any of the K1DS components had been delivered (baseline) and after 2.5 years of intervention (Spring 2015). Figure 6 Panel B shows that baseline differences in quality were quite large (e.g. $\sim 1SD$ or larger) but gaps in quality were reduced by about half or more for three out of four literacy/language and mathematics quality measures. Gaps, however, widened for math instructional quality. Figure 7 compares K1DS and BPS process quality, as measured by the CLASS in spring 2015, both in terms of mean scores (Panel A) and standardized differences (Panel B). Differences were modest for Emotional Support and large and statistically significant for Organization and Instructional Support.

Figure 6 Comparison of CBO and BPK classrooms instructional quality as measured by the ELLCO and COEMET at baseline and after 2.5 years.

Panel A: Means comparison



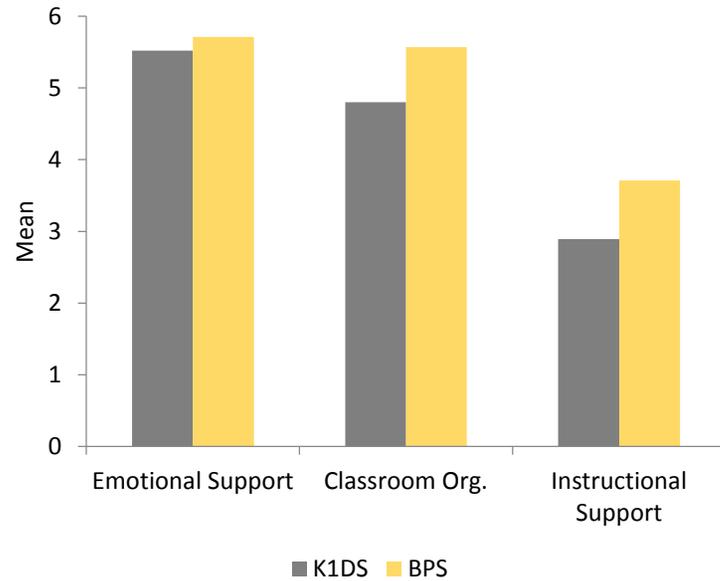
Panel B: Standardized differences comparison



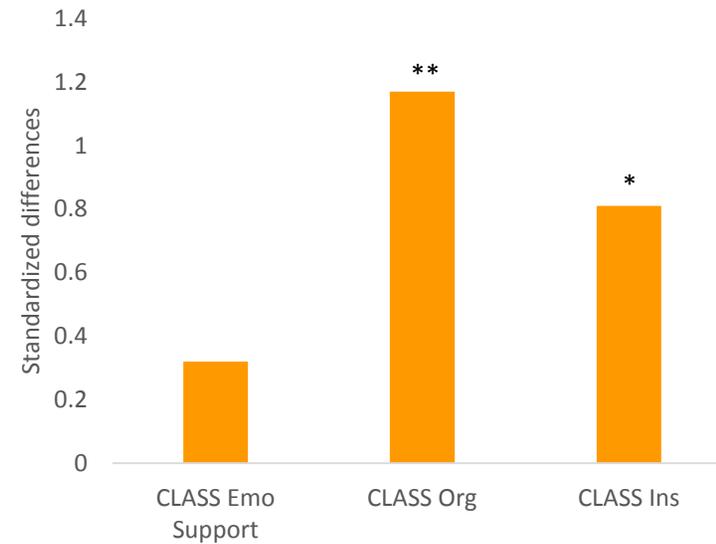
Note: CBO N=13 at baseline and N=10 in Spring 2015. BPS N=23 in Spring 2015. Panel A displays mean scores by time point and auspice only. Panel B displays standardized differences between K1DS and BPS classroom, which were computed by dividing the difference between CBO K1DS and BPS K1 scores by the BPS K1 Spring 2015 standard deviation of the relevant subscale. Differences between CBO K1DS at baseline and BPS K1 classrooms were statistically significantly different for Language and Literacy Instructional Quality and Math Instructional. Quality ($p < .05$). Other mean differences shown in the figure were not statistically significant.

Figure 7 Average classroom instructional quality in K1DS and BPS in Spring 2015, as measured by the CLASS.

Panel A: Means comparison



Panel B: Standardized differences comparison



Note: CBO N=10 in Spring 2015. BPS N=23 in Spring 2015. Standardized differences were computed by dividing the difference between CBO K1DS and BPS K1 scores by the BPS K1 Spring 2015 standard deviation of the relevant subscale. Differences between CBO K1DS and BPS classrooms were statistically significantly different in Spring 2015 on CLASS Organizational Support and CLASS Instructional Support ($p < .05$). Other mean differences shown in the figure were not statistically significant.

Children attending K1DS classrooms showed gains in their school readiness skills.

In this exploratory pilot study, child assessment data were meant to ground our understanding of child learning in community-based K1DS classrooms and were not meant to provide information about the success (or lack thereof) of the CBO K1DS project. This is particularly the case because we lacked a matched comparison group from the district and were not able to randomize children to the CBO K1DS classrooms. Furthermore, the mixed-age groups present in 12 of the 14 classrooms means that the majority of the CBO K1DS children experienced OWL and Building Blocks implementation in a way that was potentially different from children in classrooms in which only 4-year-olds were in attendance.

Overall, children's performance on direct assessments of their receptive language, literacy, mathematics, and impulse control improved from the start of their prekindergarten year to the end of that same year. In Table 11, we present a summary of student performance on child assessments from CBO K1DS classrooms. In this table, we have averaged child assessment scores across both academic years. Cohort 1 contained 100 4-year-old children who attended CBO K1DS classrooms in the 2013—2014 school year. Cohort 2 contained 63 4-year-old children who attended CBO K1DS classrooms in the 2014-2015 school year. Given the small sample size of each cohort and the relatively stable level of classroom quality across both years, we averaged the child scores across both cohorts before comparing the performance of children from CBO K1DS classrooms with the performance of children from BPS K1 classrooms. In column 4 we list the effect sizes of a quasi-experimental study of the impact of attending a BPS K1 classroom. The effect sizes represent the standardized difference in scores between children who attended K1 and those who did not. For details about this study, please see: (Weiland & Yoshikawa, 2013). In columns 5, 6, and 7 we describe the performance of a subset of children who attended BPS K1 classrooms in the 2009-2010 school year. These children were only assessed on a subset of the instruments used in this study; therefore, there are blank rows in these columns. Where possible, we compare the scores of CBO K1DS children to children represented in columns 5-7.

It is important to remember that this study did not have an experimental design and that there are many unobserved ways in which the CBO K1DS children differed from the As the

last column of this table illustrates, children who attended CBO K1DS classrooms scored lower than children who attended BPS K1 classrooms (represented by a positive value in this column) in all but one domain. CBO K1DS children performed as well or better, on average, on an assessment of their receptive language. We provide further details about child assessment performance in the Appendix.

Table 11 Comparisons of CBO K1DS and BPS K1 child performance on a range of child assessments.

	<i>CBO K1DS Cohorts 1 & 2</i>		<i>BPS impacts</i>		<i>BPS 2009-2010</i>		<i>CBO-BPS differences</i>	
	<i>Fall</i>	<i>Spring</i>	<i>Stand</i>	<i>ES</i>	<i>Fall</i>	<i>Spring</i>	<i>Stand</i>	
	<i>Mean</i>	<i>Mean</i>	<i>difs</i>	<i>(SD)</i>	<i>Mean</i>	<i>Mean</i>	<i>difs</i>	
Receptive language	88.13 (16.53)	95.98 (13.15)	0.45	0.44	88.19 (17.63)	94.45 (17.89)	0.36	-0.09
Early reading	9.13 (5.21)	10.67 (7.41)	0.28	0.62	--	--	--	0.34
Early math	10.76 (5.34)	12.36 (7.29)	0.34	0.59	--	--	--	0.25
Working memory	4.07 (0.94)	4.41 (0.91)	0.26	0.24	3.86 (1.31)	4.46 (1.18)	0.46	0.20
Inhibitory control	7.69 (5.33)	10.42 (5.02)	0.42	0.21	8.69 (6.47)	12.94 (4.56)	0.66	0.24

Note: CBO KIDS 2013-2014: $N_{\text{children}}=100$. CBO KIDS 2013-2014: $N_{\text{children}}=63$. BPS K1 $N_{\text{children}}=2,018$. Standardized differences for the receptive language, working memory, and inhibitory control measures are reported in terms of the BPS 2009-2010 Preparing to Succeed follow-up study. Standardized differences for all other measures are reported in terms of the Preparing to Succeed control group standard deviations for each measure. The BPS sample was from the Preparing to Succeed study which included children who attended the BPS K1 Program in 2008-2009 or 2009-2010 (see Weiland & Yoshikawa, 2013). Effect sizes are listed as were reported in (Weiland & Yoshikawa, 2013) and are also reported in terms of the control group standard deviations. CBO-BPS differences compare the ES of the CBO classrooms on average to the Preparing to Succeed 2009-2010 ES.

Nine Barriers to Implementation

As the fidelity and quality data illustrate, there were some successes in bringing the K1 model to CBOs but initial gains were not fully sustained through the end of the project. Qualitative and survey data point to nine barriers to implementation.

First, in the period between January 2013 and June 2014, the majority of K1DS teachers reported challenges in planning for two curricula ($N=25$ out of 34 teachers, representing teachers in 10 out of 14 classrooms). These teachers did not use the BPS K1 pacing and planning guide and therefore moved too quickly through the OWL units compared to the expectations set forth by the guides. These struggles persisted for teachers in 5 of the 14 CBO K1DS classrooms in the 2014-2015 school year ($N=10$ out of 30). Teachers reported that the lack of planning time was particularly challenging because the new curricula conflicted with standard practices in the K1DS centers. Teachers reported that this conflict decreased the utility of administrator feedback on classroom observations as center directors or instructional support personnel (where available) were better versed in the expectations associated with standard center practice rather than with the two new curricula. Teachers also reported that this conflict was apparent as other classrooms within their center followed a different schedule and shared resources which the K1DS classrooms were no able to take advantage of.

Second, implementation was particularly challenging in classrooms where teachers did not stop using part or the entire curriculum in use prior to the CBO K1DS program. This was most evident in the 10 classrooms using Creative Curriculum prior to participating in K1DS. The expectation of this pilot program was that teachers would cease using these curricula, but some teachers persisted using previous curricula in part because many of the classrooms in these centers were also using Teaching Strategies GOLD as a child assessment and parent communication tool. Teaching Strategies GOLD has a reporting structure that constrains how teachers indicate their plans and learning activities. It also creates a report after teachers have entered data that suggests particular activities and strategies derived from the Creative Curriculum system. Both of these aspects of Teaching Strategies GOLD disrupted the full implementation of the K1DS curricula.

Third, schedules in CBO K1DS classrooms posed problems for children's exposure to intentional instruction. Teachers in 12 of the 14 classrooms struggled to implement more than 2 hours of planned, intentional instruction in the 10-hour day. This seemed due in part to the fact that 1) children arrived to the classroom in a staggered fashion, 2) the lead teacher was present for only a portion of the time when all children were present due to how work schedules were designed. The CBO centers offered an early shift—starting at 7:30 am and ending at 3:00 pm or the late shift—starting at 10:00 am and ending at 6:00 p. As a result, lead teachers missed part of the time when instruction could happen, and 3) CBO K1DS centers scheduled nap times and two meals into their days. Additionally, some teachers indicated that they took breaks from the K1DS curricula especially around holidays or during times when many children were out for vacation. Overall, the more flexible expectations in CBO K1DS centers with the start and end of the school day and the number of school days attended appears to have led to reduced curricular dosage and adherence. We expected the opposite – that a longer CBO day would result in increased instructional time.

Fourth, lack of common planning time for K1DS teachers has also presented a barrier to change. Only one of the 10 K1DS centers provided reliable coverage for consistent out-of-classroom planning time. For teachers in the other nine centers, this meant that planning time often occurred during the children's nap time. Additionally, teachers from centers without regular planning time reported taking turns to plan on weekends or after work and communicating with the other teaching team members through notes or brief conversations during the children's free choice time. Twenty-four (spring 2014) to 27 (spring 2015) teachers reported on this challenge.

Fifth, nine out of 10 center directors reported that finding and paying for predictable coverage for teachers was an on-going issue in both the 2013-2014 and 2014-2015 school years, despite K1DS providing funding for substitute coverage. This issue interfered with centers' ability to schedule coaching sessions and to institute a regular accountability system that included classroom observations and teacher-director debriefs. Relatedly, center directors reported challenges in recruiting qualified teachers (particularly those with a bachelor's degree) in light of the compensation discrepancy between average CBO pay and BPS pay. This was at

the heart of two centers' struggle to replace CBO K1DS teachers who had left prior to the 2014-2015 school year. In one of these centers, few qualified candidates applied. In the other center, a replacement teacher left mid-year after being offered a higher-paying job.

Sixth, five out of 10 centers had only one full-time administrator. Eight out of 10 center leaders reported not having sufficient time to regularly observe classrooms and provide instructional support and professional development focused on improving classroom quality. Relatedly, only two center directors reporting having a systematized accountability system which provided clear goals and expectations for teacher performance. This issue was not remedied in the second year of the program. Teacher interview data supported these findings. Twenty teachers, representing 11 classrooms, reported rarely or never receiving feedback from a center leader regarding classroom practice.

Seventh, four teachers in two of the 14 classrooms reported struggling more with challenging child behaviors in 2014-2015 compared to the prior year. These teachers did not feel adequately supported in meeting the needs of these children. Teachers felt that managing these challenging behaviors undermined fidelity to the K1 model.

Eighth, most CBO K1DS classrooms were mixed-age. Twelve classrooms enrolled both 3's and 4's and the average percentage of 3's in these classrooms was 66% (2013-2014) to 60% (2014-2015) (range of 12 to 100 in both years). This is important because the OWL and Building Blocks curricula were developed for four year olds. The K1DS coach reported that teachers struggled to differentiate the curricula adequately for both age groups, which undermined curricular implementation and instructional quality.

Finally, gains in quality may have been hard to sustain through the end of the project due to a relatively high rate of turnover (50% rate for teachers from baseline to spring 2015, described in more detail below). This turnover reduced the capacity of classrooms to implement the K1DS program components because: 1) many centers struggled to replace teachers, 2) once hired, new staff did not receive formal training on the curricula used in K1DS, and 3) collaboration within teaching teams was destabilized by the turnover.

Teacher Retention

As introduced above, teacher retention was hypothesized as a mechanism for generating desired long-term improvement in child outcomes. After each academic year, CBO K1DS classrooms experienced some turnover. We attempted to determine if the teacher turnover rates during the CBO K1DS program were different than those prior to the program; however, these data were not consistently gathered across centers. Furthermore, it was difficult to determine how centers distinguished between teacher turnover compared to teacher transition into different classrooms within the same center. Nevertheless, the most recent data from a nationally representative sample of early childhood care and education centers reported on average and across auspice type a 13% teacher turnover rate in 2012 (Whitebook, Phillips, Howe, 2014). Table 12 summarizes the three waves of teacher and director turnover that occurred during the CBO K1DS project and shows that following each of the two full academic years during which the CBO classrooms participated in the K1DS program there was a 12% turnover rate.

Teacher retention was a central component of the CBO K1DS intervention model because the three full group training sessions that focused on the content and strategies of the curricula used in the K1DS program were offered at the beginning of the intervention (the summer of 2013). Thus, teachers who participated in these training sessions were hypothesized to have a foundation of knowledge upon which the coaching sessions could build classroom strategies and implementation skills. Teachers new to the CBO K1DS classrooms after K1DS program onset did not receive such intense trainings—some received no formal introduction to the curricula. Even when they did receive training prior to entering the classrooms, five of seven teachers who entered the CBO K1DS classrooms after the summer of 2013 reported being overwhelmed by the triple challenge of orientation to a new classroom, new teaching team, and new curricula.

While exit interviews were not conducted with all CBO K1Ds teachers who left their classrooms during the study period, four of the teachers who left during the study period reported that a variety of factors including opportunities for better compensation led to their decisions to leave (personal communication with CBO K1DS teachers, various dates). The salary

supplementation offered through the CBO K1DS program differed across centers due to the role that center directors originally took in requesting supplementation amounts.

Table 12 Teacher and director turnover over the course of the CBO K1DS intervention.

	Teachers N (%)	Directors N (%)
Originally assigned to CBO K1DS classroom, left prior to 2013-2014 school year	2 (6%)	2 (20%)
Left after 2013-2014 school year	4 (12.5%)	2 (20%)
Left after 2014-2015 school year	4 (12.5%)	0 (0%)
Overall percentage of original cohort remaining in CBO K1DS classrooms at beginning of 2015-2016 school year	50%	60%

Note: Each of the ten centers employed one director. Classrooms differed in how many teachers were assigned to each classroom. Overall, there were 34 lead and assistant teachers assigned to CBO classrooms.

CBO classrooms offer unique strengths.

Despite implementation challenges, we learned from BPS K1 coach, teacher, and center director interviews that CBO K1DS classrooms offered several notable strengths as a preschool delivery setting. For example, teachers and students in most centers had family-style meals and some teachers took advantage of this time to engage children in rich conversations characterized by multiple turn-taking and that drew on student interests and experiences. These kinds of conversations build important child language skills and help prepare children for elementary school. In addition, because most programs do not provide transportation, a caregiver for each child was usually in the classroom once or twice a day. Conversations between teachers and caregivers at drop-off and pickup facilitated closer family-school relationships and family engagement. BPS K1 teachers, conversely, do not generally eat lunch with children and many children arrive at school by bus. Finally, the CBO K1DS school day was longer than the BPS K1 school day by an average of 3.5 hours. This means that in theory, K1DS teachers had more time to deliver the language, literacy, and mathematics curricula than did BPS K1 teachers.

Putting it all together: How do Fidelity, Quality, and Child Performance relate to one another?

As a way to understand how the data gathered at different levels of this project fit together, we conducted a series of bivariate correlations to determine if key elements of the project were associated with one another. In Table 13, we display statistics summarizing a

series of bivariate correlations between classroom fidelity to the implementation plan and classroom instructional quality. We found evidence of a statistically significant, positive relationship between Building Blocks fidelity and Classroom Organization as measure by the CLASS ($r = 0.83, p = 0.01$). Both Building Blocks fidelity and OWL fidelity were also positively associated with the other aspects of classroom quality measured in this study but we lacked sufficient power to determine whether these associations were statistically significant. Nevertheless, the associations were in the hypothesized direction such that classrooms with higher fidelity of implementation of either OWL or Building Blocks would also receive higher ratings of classroom quality in each of the dimensions that we investigated.

Table 13 Bivariate correlations between fidelity ratings as observed in May-June 2015 classroom instructional quality ratings in May-June 2015

	OWL fidelity	p-value	Building Blocks fidelity	p-value
Language and Literacy Instructional Quality	0.43	0.21	0.66	~ 0.08
General Quality (ELLCO)	0.40	0.25	0.62	0.10
Mathematics Instructional Quality	0.03	0.94	0.56	0.15
Mathematics Classroom Culture	0.18	0.63	0.24	0.56
Emotional Support	0.25	0.49	0.69	~ 0.06
Classroom Organization	0.53	0.12	0.83	* 0.01
Instructional Support	0.30	0.41	0.55	0.16

~ $p < 0.10$; * $p < 0.05$

We explored the bivariate correlations between classroom quality and implementation fidelity and four of the nine barriers we were able to operationalize quantitatively. In Table 14, we present evidence that there are statistically significant associations between the barriers we explored and aspects of classroom quality and Building Blocks or OWL fidelity. We found that stability within the teaching team (operationalized as having at least one CBO K1DS teacher who remained in his or her assigned classroom throughout the course of the intervention) had a positive and statistically significant correlation with language and literacy instructional quality and Building Blocks fidelity. We also found positive, statistically significant relationships between the presence of a stable director and multiple dimensions of classroom quality as measured by the ELLCO and CLASS. The percentage of 3-year-old children was correlated with lower mathematics instructional quality, lower classroom organization, and lower Building

Blocks fidelity. Finally, we found evidence of statistically significant, negative associations between the retention of the Creative Curriculum and mathematics classroom culture, classroom organization, and Building Blocks fidelity. In general, the percentage of 3-year-old children in the classroom and the retention of Creative Curriculum, the most commonly used practice in CBO K1DS classrooms prior to the intervention, was negatively correlated with all dimensions of classroom quality we investigate and fidelity to the two curricula.

Table 14 Bivariate correlations between classroom instructional quality measured in May-June 2015 and the stability of the CBO K1DS center director and teaching staff.

	Stable teaching team		Stable director		% 3-yr-olds		Creative Curriculum †	
Language and Literacy Instructional Quality	0.73	*	0.77	**	-0.31		-0.46	
General Quality (ELLCO)	0.61	~	0.77	**	-0.35		-0.53	
Mathematics Instructional Quality	0.00		0.27		-0.79	*	-0.56	
Mathematics Classroom Culture	0.19		0.27		-0.31		-0.67	*
Emotional Support	0.44		0.82	**	-0.46		-0.61	~
Classroom Organization	0.19		0.58	~	-0.70	*	-0.71	*
Instructional Support	0.55	~	0.72	*	-0.08		-0.42	
OWL fidelity	0.50	~	0.32		-0.33		-0.52	
Building Blocks fidelity	0.61	*	0.67	*	-0.72	*	-0.71	*

~p < 0.10; *p < 0.05; †= Classroom retained Creative Curriculum in addition to Building Blocks and OWL.

Conclusions

In our evaluation of the K1DS project, we investigated a range of child-, teacher-, and center-level phenomena. In all, the story these data tell is one of potential within the CBO K1DS centers for quality improvement. Yet this potential is threatened by structural challenges, some of which the K1DS theory of change anticipated and attempted to address and others which were unexpected. The BPS K1 model can be implemented in community-based settings but there are several issues which must be anticipated and planned for in order for classroom quality to improve in a systematic and sustained fashion.

In this report, we discussed challenges that centers faced in implementing the two curricula that lie at the center of the BPS K1 model. Together, these challenges characterize preschool and prekindergarten settings in which classroom quality improvements are fragile and sensitive to shifts in classroom and center staffing as well as child enrollment. This point is

underscored by the bivariate correlations which show that fidelity to the implementation of the curricula is positively associated with stability within the classroom teaching team and center leadership. Furthermore, stable center leadership was positively correlated with many aspects of high quality classrooms. These connections suggest that structural aspects of CBO prekindergarten centers and classrooms may have to be in place before an intervention such as the K1DS program is put into place. Given the findings from this study, we put forward the following recommendations.

Recommendations

Policy

- **Classroom quality, particularly instructional quality can be systematically improved in community-based prekindergarten programs.**
 - **Positive change in instructional quality was fragile because classroom quality in CBO K1DS classrooms was sensitive to teacher and leader turn-over, the presence of children with challenging behaviors, among other factors.**
 - **Sustainable growth in classroom quality takes time and investment. In the CBO K1DS program, programs benefitted most from targeted coaching support.**
- **A CBO center's capacity for offering a range of systematic professional supports for teachers and leaders makes a difference.**
 - **Centers with the infrastructure to provide stable instructional leadership, predictable out-of-classroom planning time, an accountability system for classroom relational and instructional practices experienced the highest, sustained levels of classroom quality.**
 - **A center's capacity varied by organization size and structure.**
- **Balance support with accountability**
 - **Provide supports to centers but also put in place a clear accountability system for implementation and performance (Duncan & Murnane, 2013). Centers generally did not have in place strong teacher and director review processes or a strong culture of accountability. Link program participation to meeting a set of clear performance standards.**

Practice and Implementation

- **Provide more interaction between K1DS teachers and master implementers of Opening the World of Learning and Building Blocks.**
 - **Classroom visits or video clips offer opportunities to understand the implications of high levels of implementation of the curricula.**
- **Make clear that the curricula are the critical component of delivering the BPS K1 model and not a supplement or add-on.**
 - **Classrooms should suspend the bulk of activities associated with former practice or formerly used curricula**

- In particular, integrate clear implementation fidelity expectations into training, coaching, and program quality monitoring. Link program participation to meeting these expectations.
- **Focus implementation efforts on classrooms which enroll 4-year-old children only.**
 - When this is not attainable, restrict the number of three year olds enrolled and provide extra supports for differentiating the curricula.
- **Make training and coaching adaptive, in anticipation of higher teacher turnover**
 - Elevated rates of teacher turnover relative to BPS are likely to continue in any CBO expansion effort, due to factors such as low compensation among others. Revise the current training model to adapt to this reality, including formalizing how to bring new teachers up to speed via videotapes of training and additional resources for coaching.
- **Start small because piloting:**
 - Highlights unanticipated barriers to implementation fidelity
 - Allows program designers to adapt components to suit the capacity of target centers and classrooms
 - Provides an opportunity for participants in the program to provide input on components, pacing, and capacity
 - Reveals baseline capacity needs for any center or classroom wishing to participate in the program
- **Support the inclusion of regular, shared planning time for all teaching teams on a predictable schedule.**
 - Build a center-wide schedule that allows coverage for classrooms during rest time so that at least two teaching team members are available to plan outside the classroom. Some centers have attained this by staffing all classrooms with three teachers.
 - Relatedly, ensure that qualified substitute teachers are available and that the core teaching teams in K1DS classrooms are not regularly responsible for covering teacher absences in other classrooms.
- **Take steps to support stable teaching teams and leadership.**
 - Living wages and benefits packages may reduce turnover
 - Professionalization of the early childhood workforce may be attained through better instructional and classroom management support by center leadership.
 - Relatedly, support staff may help reduce center leadership administrative duties freeing leaders to provide professional support to teachers.
 - Create a collaborative space for community-based K1DS center leadership that offers specific, actionable support for data-driven instructional leadership and task management.

Research

- **Design an evaluation plan that is both summative and formative**
 - Offer real-time, actionable findings to key stakeholders so that the program model can adapt to the realities discovered during implementation.

- **Mixed methods approaches allow all participants a chance to share their unique perspective on the program.**
- **Researcher-practitioner partnerships reduce “translation” gaps by forcing stakeholders to develop and use a shared understanding of program goals, practitioner challenges, and the trade-offs involved in educational change.**

References

- Ansari, A., Purcell, K., & Gershoff, E. (2015). Classroom age composition and the school readiness of 3- and 4-year-olds in the Head Start program. *Psychological Science, 27*(1), 1-11.
- Ashe, M. K., Reed, S., Dickinson, D. K., Morse, A. B., & Wilson, S. J. (2009). Opening the World of Learning: Features, effectiveness, and implementation strategies. *Early Childhood Services, 3*, 179–191.
- Barnett, W.S., Carolan, M.E., Squires, J.H., Clarke Brown, K., & Horowitz, M. (2015). The state of preschool 2014: State preschool yearbook. New Brunswick, NJ: National Institute for Early Education Research.
- Bierman, K. L., Nix, R. L., Greenberg, M. T., Blair, C., & Domitrovich, C. E. (2008). Executive functions and school readiness intervention: Impact, moderation, and mediation in the Head Start REDI program. *Development and Psychopathology, 20*, 821–843.
- Center on the Developing Child (2010). The Foundations of Lifelong Health Are Built in Early Childhood. Retrieved from www.developingchild.harvard.edu.
- Clements, D. H., & Sarama, J. (2007a). *SRA real math, PreK-building blocks*. Columbus, OH: SRA/McGraw-Hill.
- Clements, D. H., & Sarama, J. (2007b). Effects of a preschool mathematics curriculum: Summative research on the Building Blocks project. *Journal for Research in Mathematics Education, 38*, 136–163.
- Diamond, A., & Taylor, C. (1996). Development of an aspect of executive control: Development of the abilities to remember what I said and to “do as I say, not as I do.” *Developmental Psychobiology, 29*, 315–344.
- Dickinson, D. K., Freiberg, J. B., & Barnes, E. (2011). Why are so few interventions really effective? A call for fine-grained research methodology. In S. B. Neuman & D. K. Dickinson (Eds.), *Handbook of early literacy research* (Vol. III, pp. 337–357). New York: Guilford Press.
- Dickinson, D. K., Kaiser, A., Roberts, M., Hofer, K. G., Darrow, C. L., & Griffenhagen, J. B. (2011). *The effects of two language focused preschool curricula on children’s achievement through first grade*. Paper presented at the Society for Research in Educational Effectiveness, Washington, DC.
- Duncan, G.J. & Magnuson, K. (2011). The nature and impact of early achievement skills, attention skills, and behavior problems. In G.J. Duncan and R.J. Murnane (eds.), *Whither*

- Opportunity: Rising Inequality, Schools, and Children's Life Chances (pp.47-70). New York: Russell Sage.
- Dunn, L. M., & Dunn, L. M. (1997). *Peabody picture vocabulary test-third edition*. Bloomington, MN: Pearson Assessments.
- Gathercole, S. E., & Pickering, S. J. (2000). Working memory deficits in children with low achievements in the national curriculum at 7 years of age. *British Journal of Education Psychology, 70*, 177–194.
- Gormley, W. T., Gayer, T., Phillips, D., & Dawson, B. (2005). The effects of universal pre-K on cognitive development. *Developmental Psychology, 41*, 872–884.
- Gormley, W. T., Phillips, D. A., & Gayer, T. (2008). Preschool programs can boost school readiness. *Science, 320*, 1723–1724.
- Jan, T. (2007). “Boston Preschools Falling Far Short of Goals, Study Says.” Boston Globe. Section A, April 7, 2007.
- Lin, R.Y. (2009). *Case Study Research Design and Methods*. Thousand Oaks, CA: Sage Publishing.
- Maxwell, J.A. (2012). *Qualitative Research Design: An Interactive Approach*. Thousand Oaks, CA: Sage Publishing.
- Peisner-Feinberg, E. S., Burchinal, M. R., Clifford, R. M., Culkin, M. L., Howes, C., Kagan, S. L., et al. (2001). The relation of preschool child-care quality to children’s cognitive and social development trajectories through second grade. *Child Development, 72*, 1534–1553.
- Pianta, R. C., La Paro, K. M., & Hamre, B. K. (2008). *Classroom assessment scoring system [CLASS] manual: Pre-K*. Baltimore, MD: Brookes Publishing.
- Schickedanz, J., & D. Dickinson. (2005). *Opening the world of learning*. Iowa City, IA: Pearson Publishing.
- Sarama, J. & Clements, D.H. (2009). *Manual for Classroom Observation of Early Mathematics—Environment and Teaching*. Unpublished manuscript.
- Sarama, J., Lange, A. A., Clements, D. H., & Wolfe, C. B. (2012). The impacts of an early mathematics curriculum on oral language and literacy. *Early Childhood Research Quarterly, 27*, 489-502.
- Smith, M. W., Dickinson, D. K., Sangeorge, A., & Anastasopoulos, L. (2002). *User’s guide to the Early Language & Literacy Classroom Observation toolkit: Research edition*. Baltimore, MD: Brookes.

- Smith-Donald, R., Raver, C. C., Hayes, T., & Richardson, B. (2007). Preliminary construct and concurrent validity of the Preschool Self-Regulation Assessment (PSRA) for field-based research. *Early Childhood Research Quarterly, 22*, 173–187.
- U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse. (2013, March). Early Childhood Education intervention report: The Creative Curriculum® for Preschool, Fourth Edition. Retrieved from <http://whatworks.ed.gov>.
- Wanless, S.B. (2013). *An assessment of a (pre)school's readiness for change*. University of Pittsburgh, Unpublished measure.
- Wechsler, D. (1986). *Wechsler Intelligence Scale for Children—Revised*. New York: Psychological Corporation.
- Weiland, C. (2016a). Inclusion preschool in the Boston Public Schools: Impacts and dosage. Manuscript under review.
- Weiland, C. (2016b). Preschool 2.0: High-quality public preschool programs at scale. Manuscript under review.
- 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. doi:10.1016/j.ecresq.2012.12.002
- 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.
- Whitebook, M., Phillips, D., & Howes, C. (2014). *Worthy work, STILL unlivable wages: The early childhood workforce 25 years after the National Child Care Staffing Study*. Berkeley, CA: Center for the Study of Child Care Employment, University of California, Berkeley.
- Wilson, S. J., Morse, A. B., & Dickinson, D. K. (2009). *Examining the effectiveness of OWL as used in ERF Projects: Final report of results from the OWL Consortium Project*. Nashville, TN: Vanderbilt University Center for Evaluation Research & Methodology.
- Wong, V. C., Cook, T. D., Barnett, W. S., & Jung, K. (2008). An effectiveness-based evaluation of five state prekindergarten programs. *Journal of Policy Analysis and Management, 27*, 122–154.
- Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson Tests of Achievement*. Itasca, IL: Riverside.
- Yoshikawa, H., Weiland, C., Brooks-Gunn, J., Burchinal, M. R., Espinosa, L. M., Gormley, W. T., Ludwig, J., Magnuson, K. A., ... & Zaslow, M. J. (2013). Investing in our future: The

evidence base on preschool education. Society for Research in Child Development and Foundation for Child Development.

Appendix

Table 1 Instruments used in CBO K1DS evaluation project

Construct	Measure	Embedded In	Validity and Reliability	Citation
<i>Child Measures</i>				
Receptive Vocabulary	Peabody Picture Vocabulary Test III (PPVT-III)	Fall and Spring Child Direct Assessments	Test-Retest: $r=0.85-0.90$ Internal Consistency: $\alpha=0.95$ Concurrent/ Predictive Validity: other measures of language, literacy and academic performance	(Dunn & Dunn, 1997)
Literacy and Pre-reading Skills	Woodcock–Johnson Letter-Word Identification subscale (WJ-LW)	Fall and Spring Child Direct Assessments	Test-Retest: $r=0.96$ Concurrent/ Predictive Validity: other measures of language and literacy and academic performance	(Woodcock, McGrew, & Mather, 2001)
Numeracy, Early Mathematics and Problem-Solving Skills	Woodcock–Johnson Applied Problems subscale (WJ-AP)	Fall and Spring Child Direct Assessments	Test-Retest: $r=0.92-0.94$ Internal Consistency: $\alpha=0.85$ Concurrent/ Predictive Validity: other measures of early mathematics and academic performance	(Woodcock et al., 2001)
Inhibitory control	Pencil Tap	Fall and Spring Child Direct Assessment		(Diamond & Taylor, 1996)
Working Memory	Forward Digit Span (FDS)	Fall and Spring Child Direct Assessment		(Gathercole & Pickering, 2000)
Attention Shifting, Positive	Task Orientation	Fall and Spring Child		(Smith-Donald, Raver, Hayes,

Emotion, Impulse Control	Questionnaire (TOQ)	Direct Assessment		& Richardson, 2007)
<i>Classroom Measures</i>				
Quality of Interactions	COEMET, CLASS & ELLCO	Classroom Observations, Baseline and Spring follow-ups	See below	
Mathematics Instruction/Use of Building Blocks Curriculum	Classroom Observation of Mathematics—Environment and Teaching (COEMET)	Classroom Observations, Baseline and Spring follow-ups	Internal Consistency: $\alpha_{total}=0.96$	(Samara & Clements, 2009)
Language Instruction/ Use of Opening the World of Learning Curriculum	Early Language and Literacy Classroom Observation tool (ELLCO)	Classroom Observations, Baseline and Spring follow-ups	Internal Consistency: $\alpha_{total}=0.84$ Construct Validity: three subscales replicated over multiple studies Concurrent/Predictive Validity: receptive vocabulary, early literacy skills	(Smith et al., 2002)
Process Quality	Classroom Assessment Scoring System	Classroom Observations, Baseline and Spring follow-ups		(Pianta, La Paro, & Hamre, 2008)
Fidelity of Implementation	Building Blocks Near Fidelity Tool and OWL fidelity tool	Classroom Observations, Baseline and Spring follow-ups		(Sarama, et al., 2012) and (Weiland & Yoshikawa, 2013)
<i>Teacher and Director Survey Measures</i>				
Readiness for Change	An Assessment of a (Pre)School's Readiness for Change	Fall Coach, Teacher, and Director Survey	Under development	(Wanless, 2013)

Table 2 CBO K1DS quality scores

	<i>All available valid data at each time pt</i>				<i>Classrooms measured at all time pts (N=9)</i>			
	Mean (SD)	Min	Max	N (%) classrooms meeting benchmark	Mean (SD)	Min	Max	N (%) classrooms meeting benchmark
Baseline (N=13)								
COEMET								
Classroom culture	3.08 (0.75)	1.67	4.11	5 (38%)	3.33 (0.89)	1.83	4.50	5 (56%)
SMA average	3.00 (0.40)	2.30	3.50	1 (10%)	3.01 (0.40)	2.26	3.51	1 (11%)
N SMA	2.40 (1.70)	1	7	--	2.78 (1.86)	1	7	--
Length SMA	6.51 (3.13)	3	13	--	6.24 (3.28)	3	13	--
N Mini SMA	2.90 (1.7)	0	6	--	2.89 (1.90)	0	6	--
ELLCO								
Language and Literacy	2.76 (1.01)	1.38	3.88	6 (46%)	2.63 (0.99)	1.50	3.88	3 (33%)
Environment	3.35 (0.88)	1.83	4.67	5 (38%)	3.50 (0.87)	2.33	4.67	4 (44%)
Spring 2014 (N=14)								
COEMET								
Classroom culture	3.85 (0.51)	3.13	4.67	9 (64%)	3.72 (0.53)	3.13	4.56	5 (56%)
SMA average	3.02 (1.32)	0	3.89	8 (57%)	3.22 (1.23)	0	3.83	6 (67%)
N SMA	2.29 (2.64)	0	10	--	2 (1.66)	0	6	--
Length SMA	8.75 (7.95)	0	33	--	10.94 (9.14)	0	33	--
N Mini SMA	2.29 (1.77)	0	6	--	2.22 (1.72)	0	5	--
ELLCO								
Language and Literacy	3.87 (0.49)	2.75	4.63	11 (79%)	3.94 (0.32)	3.38	4.38	8 (89%)
Environment	4.12 (0.58)	3.17	4.83	12 (86%)	4.20 (0.56)	3.17	4.83	8 (89%)
Spring 2015 (N=10)								

COEMET								
Classroom culture	3.61 (0.57)	2.63	4.44	7 (70%)	3.52 (0.52)	2.63	4.11	6 (67%)
SMA average	2.83 (1.57)	0.00	4.12	6 (60%)	2.71 (1.62)	0	4.12	5 (56%)
N SMA	1.90 (1.20)	0.00	3.00	--	1.78 (1.20)	0	3	--
Length SMA	6.20 (4.42)	0.00	13.33	--	6.30 (4.67)	0	13.33	--
N Mini SMA	2.60 (1.71)	0.00	5.00	--	2.44 (1.74)	0	5	--
ELLCO								
Language and Literacy Environment	3.10 (0.72)	2.33	4.75	2 (20%)	2.92 (0.45)	2.33	3.50	1 (11%)
Environment	3.60 (0.51)	2.86	4.71	6 (60%)	3.48 (0.35)	2.86	3.86	5 (56%)
CLASS								
Emotional Support	5.52 (0.63)	4.60	6.40	7 (70%)	5.42 (0.58)	4.60	6.10	6 (67%)
Classroom Org.	4.80 (0.85)	3.67	6.27	5 (50%)	4.70 (0.83)	3.67	6.27	4 (44%)
Instructional Support	2.89 (1.06)	1.60	5.20	1 (10%)	2.63 (0.72)	1.60	3.67	0 (0%)

Note: Benchmarks were set as follows: 3.5 for ELLCO and COEMET; 4.5 for CLASS Instructional Support; 5 for CLASS Emotional Support and Classroom Organization.

Table 3 BPS K1 quality scores

				% classrooms meeting benchmark
<i>COEMET (N=10),</i>				
Winter 2013				
Classroom culture	4.10 (0.36)	3.56	4.56	10 (100%)
SMA average	3.71 (0.39)	3.05	4.52	9 (90%)
N SMA	3.30 (0.82)	2	4	--
Length SMA	11.81 (7.10)	5	27	--
N Mini SMA	3.70 (2.75)	1	10	--
<i>ELLCO (N=9)</i>				
Language and Literacy	4.04 (0.86)	2.5	5.0	7 (70%)
Environment	3.94 (1.05)	2.33	5.0	8 (80%)
.....				
Spring 2015 (N=23)				
<i>COEMET</i>				
Classroom culture	3.61 (0.44)	2.67	4.38	13 (57%)
SMA average	3.51 (0.85)	0.00	4.37	16 (70%)
N SMA	3.22 (2.04)	0.00	8.00	--
Length SMA	9.01 (4.78)	0.00	20.00	--
N Mini SMA	2.57 (2.41)	0.00	10.00	--
<i>ELLCO</i>				
Language and Literacy	3.44 (0.61)	2.33	4.67	14 (61%)
Environment	3.77 (0.50)	2.71	4.71	17 (74%)
<i>CLASS</i>				
Emotional Support	5.71 (0.60)	4.70	6.55	22 (96%)
Classroom Org.	5.57 (0.66)	4.27	7.00	22 (96%)
Instructional Support	3.71 (1.01)			4 (17%)
		2.20	5.73	
.....				
Spring 2010 (N=83)				
<i>CLASS</i>				
Emotional Support	5.63 (0.60)	4.00	6.83	73 (88%)
Classroom Org.	5.10 (0.68)	2.75	6.22	56 (67%)
Instructional Support	4.30 (0.84)	2.22	5.67	42 (51%)
<i>ELLCO</i>				
Language and Literacy	3.53 (0.45)	2.50	4.50	48 (58%)
Environment	3.79 (0.72)	2.20	5.00	51 (61%)

Note: Benchmarks were set as follows: 3.5 for ELLCO and COEMET; 4.5 for CLASS Instructional Support; 5 for CLASS Emotional Support and Classroom Organization.

Table 4 Characteristics of 4 year olds in the full sample compared to those who were assessed.

	2013-2014		2014-2015	
	Full Sample	Children Assessed in Fall and Spring	Full Sample	Children Assessed in Fall and Spring
Percentage of Children who were...				
Latino/a	26	28	34	34
African American	57	54	49	51
White	8	9	7	6
Asian American	6	5	6	4
Native English Speakers	67	70	69	71
Receiving financial assistance to attend prekindergarten (CBOs only)	87	70	89	73

Table 5 Comparisons of CBO K1DS cohort 1 and CBO K1DS cohort 2 performance on a range of child assessments.

	<i>CBO K1DS Cohort 1</i>		Standardized Difference (SD)	<i>CBO K1DS Cohort 2</i>		Standardized Difference (SD)	Difference between Cohort 2 and Cohort 1
	Fall 2013	Spring 2014		Fall 2014	Spring 2015		
	Mean (SD)	Mean (SD)		Mean (SD)	Mean (SD)		
Receptive language	90.18 (13.81)	95.71 (11.71)	0.34	87.83 (18.44)	96.13 (15.46)	0.51	0.17
Early reading	8.49 (4.35)	12.06 (6.42)	0.75	10.19 (5.17)	12.69 (6.98)	0.53	-0.22
Early math	11.80 (3.93)	16.07 (3.52)	1.00	12.57 (4.57)	15.32 (4.48)	0.65	-0.35
Working memory	4.01 (0.92)	4.31 (0.95)	0.31	4.10 (1.00)	4.48 (0.85)	0.40	0.09
Inhibitory control	9.83 (5.80)	13.69 (3.27)	0.62	8.21 (6.62)	11.36 (5.55)	0.51	-0.11
Attention	3.19 (0.68)	3.23 (0.53)	0.05	3.24 (0.78)	3.37 (0.58)	0.18	0.13
Impulse control	3.21 (0.53)	3.31 (0.46)	0.18	3.67 (0.60)	3.57 (0.55)	-0.18	-0.36
Positive Emotion	3.48 (0.63)	3.50 (0.53)	0.03	3.52 (0.68)	3.48 (0.63)	-0.06	-0.09

Note: CBO KIDS 2013-2014: N_{children}=100. CBO KIDS 2014-2015: N_{children}=63. The receptive language scores are standardized PPVT scores while the rest of the scores reported are raw scores. Standardized differences were calculated by dividing the difference between fall and spring scores by the pooled standard deviation of the fall average scores across both cohorts.

Table 6 Parent responses to a questionnaire regarding their choice of care and education services.

	N	% of total	
<i>Why did you choose your current center?*</i>			
The center is close to my home.	15	31%	
My child attended this center as an infant.	20	42%	
The center offers the right hours for my work schedule.	23	48%	
The cost of care at this center is affordable.	30	63%	
I did not get a spot in BPS K1.	8	17%	
I applied for a K1 spot in BPS for my child.	11	23%	
			% of those who reported turning down a K1 spot
<i>If you were offered a spot but did not take it, why not?*</i>			
I wasn't offered the school I wanted for my child.	3	6%	100%
The program didn't cover enough hours.	3	6%	100%
I was offered the spot too late in the year. I worried about transition.**	1	2%	33%

Note: * All parents whose children were 4-years-old on September 1, 2014 responded to this part of the survey. These were parents who could have placed their children in BPS K1 classrooms. ** This response was a write-in response by one parent.