Early Childhood Research Quarterly 47 (2019) 218–228

Contents lists available at ScienceDirect

Early Childhood Research Quarterly

Cumulative classroom quality during pre-kindergarten and kindergarten and children's language, literacy, and mathematics skills

Robert C. Carr*, Irina L. Mokrova, Lynne Vernon-Feagans, Margaret R. Burchinal

Frank Porter Graham Child Development Institute, The University of North Carolina at Chapel Hill, United States

ARTICLE INFO

Article history:
Received 14 November 2017
Received in revised form 3 December 2018
Accepted 15 December 2018

Keywords:
Pre-K
Kindergarten
Classroom quality
Language
Literacy
Mathematics

ABSTRACT

Prior research indicates that preschool classroom quality is positively associated with children's emergent academic skills at the end of the preschool period, but these associations appear to diminish during kindergarten (K). To consider the accumulation of classroom experiences during preschool and K, data were drawn from a sample of children ($N = 1015$) who participated in state-funded pre-K programs across six states and were followed into the spring of K. We examined if pre-K and K classroom quality were associated with children's language, literacy, and mathematics skills in the spring of K after controlling for pre-K entry skills, child background demographics, and other classroom covariates. Classroom quality was assessed with the Classroom Assessment Scoring System. We found that *additive* cumulative effects described associations between classroom quality and children's language and literacy outcomes. For language skills, higher-quality pre-K Instructional Support predicted better K language skills at all levels of Instructional Support quality in K, but higher-quality K Instructional Support also contributed to the prediction of better K language skills in an additive manner. For literacy skills, higher-quality pre-K Instructional Support, Emotional Support, and Classroom Management predicted better K literacy skills at all levels of K quality, but each domain of K quality also contributed to the prediction of better K literacy skills in an additive manner. In contrast, for mathematics, we found that *multiplicative* cumulative effects described associations between classroom quality and K mathematics skills such that the positive predictive effect of pre-K Instructional Support and Emotional Support was enhanced as the level of K quality increased in each of these domains, respectively. Moreover, sensitivity analyses revealed that higher-quality pre-K Instructional Support and Emotional Support significantly predicted better K mathematics skills for children who experienced *high*, but not *average* or *low* levels of K quality in these domains, respectively. Implications are discussed in relation to classroom quality alignment during pre-K and the early elementary grades.

A large body of research has documented positive, albeit modest associations between preschool classroom quality and children's emergent academic skills during the preschool period (see Burchinal, 2018 for a review). Yet, despite these encouraging findings, studies that have followed children into elementary school generally find these associations begin to diminish during kindergarten (K) and rarely persist into later grades (Anderson & Phillips, 2017; Belsky et al., 2007; Burchinal et al., 2008; Choi, Elicker, Christ, & Dobbs-Oates, 2016; Hindman, Skibbe, Miller, & Zimmerman, 2010; Peisner-Feinberg et al., 2001; Wen, Bulotsky-Shearer, Hahs-Vaughn, & Korfmacher, 2012; Zucker, Cabell, Justice, Pentimonti, & Kaderavek, 2013). Causal evaluation and quasi-experimental studies also provide compelling evidence to indicate that preschool program participation can have positive short-term impacts on children's academic skills during the preschool period, but the literature is less conclusive concerning the longer-term impacts into elementary school (Phillips et al., 2017; Yoshikawa et al., 2013). Most of these studies have focused on the long-term impacts of preschool experience without considering children's subsequent school experiences. To address questions about the accumulation of classroom experiences during preschool and elementary school, the current study examined the extent to which kindergarten (K) classroom quality enhanced or diminished the effects of pre-K classroom quality in relation to children's academic outcomes at the end of K.

1. Classroom quality

Vandell and Wolfe (2000) have described *process quality* as the actual experiences that children have in their classroom set-

* Corresponding author at: Frank Porter Graham Child Development Institute, The University of North Carolina at Chapel Hill, CB 8180, Chapel Hill, NC 27599-8180, United States.
E-mail address: rc@unc.edu (R.C. Carr).

https://doi.org/10.1016/j.ecresq.2018.12.010

tings. The quality of interactions between teachers and children is now the most widely accepted index of early childhood classroom process quality (Burchinal, Magnuson, Powell, & Soliday Hong, 2015). Socioecological theory and empirical research suggest that the quality and frequency of teacher–child interactions provide the foundation for early development (Bronfenbrenner & Morris, 2006; Burchinal, 2018; Burchinal et al., 2016). Moreover, while children appear to benefit from attending a center-based early childhood education program, development is clearly enhanced when children's interactions with their teachers are cognitively stimulating, emotionally supportive, and well-organized (i.e., high-quality; Phillips et al., 2017; Yoshikawa et al., 2013). In the current study, we focused on classroom process quality in pre-K and K programs as assessed by the Classroom Assessment Scoring System (CLASS; Pianta, La Paro, & Hamre, 2008), which is a theoretically informed and widely used measure of classroom teacher–child interactions and organizational features.

The accumulation of higher-quality classroom experiences over multiple years of early childhood education is also expected to play an important role in early development. These cumulative effects could be *additive* such that more years of higher-quality experiences positively contribute and more years of lower-quality experiences negatively contribute. They could also be *multiplicative* such that later high-quality experiences disproportionately enhance the effect of earlier high-quality experiences and later low-quality experiences disproportionately exacerbate the effect of earlier low-quality experiences. Multiplicative classroom quality is well aligned with the cumulative model of human capital acquisition (Cunha & Heckman, 2007). In this model, the concept of *dynamic complementarity* is used to describe how investments in children's skill formation at earlier stages of development serve as a base for investments at later stages to produce the "skill begets skill" phenomenon. The concept of dynamic complementarity posits that (1) later investments in children's skill development (e.g., higher-quality K classrooms in our case) bolster the productivity of earlier investments (e.g., higher-quality pre-K classrooms); and (2) earlier investments will not be productive unless they are followed up by later investments. In the current study, we explicitly examined both additive and multiplicative cumulative effects of classroom quality in relation to children's language, literacy, and mathematics skills in the spring of kindergarten (further described in *The Current Study* section below).

1.1. Preschool classroom quality and academic skills

Many studies report small, but positive associations between classroom quality during the preschool period and children's concurrent development of early academic skills (Auger, Farkas, Burchinal, Duncan, & Vandell, 2014; Burchinal, Vandergrift, Pianta, & Mashburn, 2010; Burchinal et al., 2016; Hamre, Hatfield, Pianta, & Jamil, 2014; Howes et al., 2008; Keys et al., 2013; Mashburn et al., 2008; Moiduddin, Aikens, Tarullo, West, & Xue, 2012; NICHD Early Child Care Research Network, 2002; Peisner-Feinberg et al., 2001; Weiland, Ulvestad, Sachs, & Yoshikawa, 2013; Wen et al., 2012; Zucker et al., 2013). The few studies that have followed children into elementary school find that these quality-outcome associations begin to diminish as early as K and too often disappear during later grades (Anderson & Phillips, 2017; Belsky et al., 2007; Burchinal et al., 2008; Choi et al., 2016; Hindman et al., 2010; Peisner-Feinberg et al., 2001; Wen et al., 2012; Zucker et al., 2013). While multiple studies have examined associations between preschool classroom quality and the concurrent or long-term development of children's academic skills, few studies have examined classroom quality in both the preschool and elementary school years in relation to child outcomes.

We found only two studies that examined multiple years of classroom quality across the preschool and elementary school periods. First, a reanalysis of data from the NICHD Study of Early Child Care and Youth Development by Ansari and Pianta (2018b) examined the cumulative effects of child care and elementary school quality. The original analyses of these data found robust, long-term associations between child care quality—measured between 6- and 54-months of age—and children's academic and behavioral outcomes at the end of high school (Vandell, Burchinal, & Pierce, 2016), but the magnitude of associations was diminished during early elementary school (Belsky et al., 2007). In their reanalysis of these data, Ansari and Pianta (2018b) found an additive effect of quality in relation to ninth grade mathematics such that child care quality and elementary school quality were both independently related to ninth grade mathematics (child care + elementary school quality). Additionally, a multiplicative effect (child care × elementary school quality) indicated that child care quality was a stronger predictor of child outcomes in fifth grade (language/literacy) and ninth grade (math & language/literacy) if children subsequently experienced higher-quality classrooms in elementary school. While these studies focused on a measure of emotional support quality during early childhood and elementary school, they did not examine other dimensions of process quality likely to be more proximal to children's development of academic skills.

A second study examined the cumulative effects of classroom quality during pre-K and K in relation to children's social and behavioral, but not academic skills (Broekhuizen, Mokrova, Burchinal, Garrett-Peters, & The FLP Key Investigators, 2016). Using a birth cohort sample of children in low-wealth rural communities, Broekhuizen et al. (2016) found that children demonstrated better social skills and fewer problem behaviors at the end of K when they experienced two years of higher-quality emotional support and classroom organization during pre-K and K, while the social/behavioral skills of children who only experienced higher-quality classrooms in either pre-K or K were lower than children who experienced two years of higher-quality and not reliably different from children who experienced no years of higher-quality—providing support for the multiplicative cumulative effect of classroom quality. Analyses of children's first grade outcomes in this same study revealed that children in the two years of higher-quality group continued to outperform all other groups, except the higher-quality in pre-K only group—suggesting that higher-quality experiences in pre-K were driving the benefit of higher-quality experiences during both years.

Results of the two aforementioned studies lend support to the propositions of dynamic complementarity (Cunha & Heckman, 2007). The study by Ansari and Pianta (2018b) demonstrated that higher quality elementary school classrooms (i.e., later investments) both *added to* and *enhanced* the effects of higher-quality child care experiences—providing support for the first proposition of dynamic complementarity; that later investments bolster the productivity of earlier investments. However, Broekhuizen et al. (2016) found that experiencing a higher-quality classroom environment in pre-K was only related to child outcomes at the end of K if children also experienced higher-quality classroom environments in K—providing support for the first and second propositions of dynamic complementarity; that later investments bolster the productivity of earlier investments and earlier investments will not be productive unless they are followed up by later investments. Despite the important contributions of these studies, several gaps remain in the literature examining classroom quality during the transition from preschool to elementary school, because Ansari and Pianta (2018b) did not have data to consider the quality of K classroom environments and Broekhuizen et al. (2016) did not consider children's academic outcomes. Additionally, the quality measures used in both studies were not focused on dimensions of classroom instructional processes, which are likely more proximal to children's development of academic skills compared to emotional support and organizational processes. Our study was designed to address these gaps and provide additional evidence to support (or refute) the propositions of dynamic complementarity.

1.2. Preschool impacts and subsequent educational contexts in elementary school

In a related line of research, studies have examined variation in the longer-term impacts of preschool programming based on characteristics of the classroom context in elementary school. The results were generally mixed, indicating that preschool impacts were sustained into elementary school under both *favorable* and *unfavorable* classroom contexts. Two studies provide examples of positive preschool impacts being sustained in more favorable elementary school classroom contexts. The K follow-up study of the Head Start REDI intervention found that impacts on children's problem solving behaviors were enhanced at the end of K for children in higher-quality K classrooms, as measured by the CLASS (Bierman et al., 2014). However, this finding was not evident for a number of other behavioral and academic skill measures. Similarly, follow-up analyses of the Building Blocks preschool curricular intervention found that impacts on children's math skills were only sustained into first grade for children whose teachers experienced additional math focused professional development (Jenkins et al., 2018). This study also found that long-term impacts were sustained in smaller class sizes. Alternatively, two studies have found that preschool impacts on children's academic skills were more likely to be sustained in unfavorable classroom contexts in elementary school, such as classrooms with less emphasis on reading instruction (Bierman et al., 2014; Magnuson, Ruhm, & Waldfogel, 2007) and larger class sizes (Magnuson et al., 2007). In addition to evidence of sustained effects in both favorable and unfavorable classroom contexts, two recent studies have not found evidence of variation in the long-term impacts of preschool programming based on subsequent classroom contexts (Bassok, Gibbs, & Latham, 2018; Jenkins et al., 2018).

It is also worth noting that the persistence of preschool impacts has also been examined in relation to school-wide context in elementary school. However, we were unable to examine this important level of analysis in the current study. A K follow-up study of the Chicago School Readiness Project found that intervention impacts on children's language, literacy, and social–emotional skills were sustained until the end of K, but only among children in elementary schools with higher levels of school-wide academic achievement (Zhai, Raver, & Jones, 2012). Alternatively, positive impacts of the Head Start REDI intervention on children's problem behaviors and social competence in K were only evident in schools with lower levels of school-wide achievement (Bierman et al., 2014). Several studies have also examined the long-term impacts of preschool program participation and the intervening role of subsequent school contexts (Ansari & Pianta, 2018a; Curenton, Dong, & Shen, 2015; Jenkins et al., 2018). A study by Ansari and Pianta (2018a) found that positive effects of preschool participation on children's academic achievement were sustained into fifth grade for children who subsequently attended high- and moderate-, but not low-quality elementary schools. Alternatively, one study found that positive effects of preschool programming on children's long-term achievement were reduced after accounting for school-wide academic achievement (Curenton et al., 2015). Finally, another study found no evidence that measures of school-wide poverty and achievement moderated the long-term effects of Head Start pre-K (Jenkins et al., 2018).

2. The current study

The current study draws on data from the National Center for Early Development and Learning (NCEDL) Multi-State Study of Pre-K—a sample of children who participated in state-funded pre-K programs across six states (Early et al., 2005). The current study extends previous research based on this sample in combination with a separate, comparable sample of children in five additional states who were not followed into K. These previous studies reported that pre-K CLASS Instructional Support quality positively predicted children's language, literacy, and mathematics skills in the spring of pre-K (Burchinal et al., 2010; Howes et al., 2008; Mashburn et al., 2008), while CLASS Emotional Support quality was positively associated with children's social–emotional and, albeit less consistently, academic skills in the spring of pre-K (Mashburn et al., 2008). The current study also extends findings from a K follow-up of the NCEDL Multi-State Study of Pre-K, which reported better language and literacy, but not mathematics skills in the spring of K for children who experienced higher-quality Instructional Support in pre-K (Burchinal et al., 2008). CLASS Instructional Support in K also predicted better language and mathematics skills in the spring of K in this follow-up study.

The current study extends previous research by examining *if* and *how* the accumulation of higher-quality of classroom environments during pre-K and K was related to children's language, literacy, and mathematics skills in the spring of K—after controlling for pre-K entry skills, child background demographics, and other classroom covariates. Moreover, based on regression analyses, we tested for two types of cumulative classroom quality effects: additive and multiplicative. The additive cumulative effect of classroom quality would be demonstrated by statistically significant and positive main effects of pre-K and K classroom quality—indicating that higher-quality pre-K uniformly contributes to the prediction of K skills at all levels of K classroom quality and that higher-quality K also contributes to the prediction of K skills. We also examined the multiplicative cumulative effect of classroom quality, which would be demonstrated by a statistically significant and positive interaction between the pre-K and K quality terms. The multiplicative cumulative effect of classroom quality would indicate that higher-quality in K disproportionately enhances the effect of higher-quality in pre-K. Moreover, the multiplicative cumulative effect is well aligned with the two propositions of dynamic complementarity: (1) later investments in children's skill development bolster the productivity of earlier investments and (2) earlier investments will not be productive unless they are followed up by later investments. The statistically significant interaction term would demonstrate only the first proposition of dynamic complementarity if a statistically significant main effect of pre-K classroom quality (i.e., earlier investments) was enhanced by the predictive effect of K classroom quality (i.e., later investments). The statistically significant interaction term would demonstrate both propositions of dynamic complementarity if a non-significant main effect of pre-K classroom quality became statistically significant at higher levels of K classroom quality. Given that prior research on preschool and elementary school classroom quality has found evidence of both additive and multiplicative cumulative effects in relation to children's academic skills (Ansari & Pianta, 2018b) and multiplicative cumulative effects in relation to social–emotional skills (Broekhuizen et al., 2016), we hypothesized that the accumulation of higher-quality classroom experiences would matter in different ways for the different academic skills considered in this study (language, literacy, and mathematics), such that the additive cumulative effect would describe associations with some skills, but not others, while the multiplicative cumulative effect would describe associations with some skills, but not others.

3. Methods

3.1. Sample

The NCEDL Multi-State Study of Pre-K was initiated during the 2001–02 school year to examine the classroom experiences and developmental competencies of 4-year-old children enrolled in center-based pre-K programs across six states (Clifford, Barbarin et al., 2005). All of the pre-K programs included in this study were fully or partially funded and operated by state and/or local education agencies, while 30% of the children in this sample were enrolled in programs that also received funding and direction from the federal Head Start program. The six states included in this study were selected on the basis of having significant pre-K initiatives. In 2001, these states served 43% of all children who participated in state-funded pre-K programs in the United States and spent 42% of the funding allocated for state pre-K nationwide (Clifford, Bryant et al., 2005).

The sampling frame for this study involved random selection of early childhood education sites within four states (Illinois, Ohio, Kentucky, and Georgia) and several regions within two states (California and New York). Pre-K classrooms were then randomly selected from sites and four children were randomly selected from each classroom. A final sample of 1015 children was recruited from 240 pre-K classrooms. After their pre-K year, the study children spread to more than 800 different K classrooms.

Descriptive statistics for each analytic variable are provided in Table 1. The average age of children in the sample was 4.56 years (SD = 0.31) in the fall of pre-K and 6.04 years (SD = 0.31) in the spring of K. About half of the children in the sample were male (49%). Of the full sample, 42% of children were identified as White by their parents, 24% as African American, 25% as Latino, and 10% as Other, which included Asian American, Native American, and Multiracial children. Based on parental report of family income, 57% of children in the sample came from families categorized as below 150% of the federal poverty line at the time children entered the study, while children from families that were categorized above 150% of the federal poverty line reported a range in family income from $17,500 to $87,500 or greater. The participating children's mothers had completed, on average, 12.52 years of schooling. Children spent an average of 24 and 27 hours per week in their pre-K and K classrooms, respectively. However, there was a wide range of variability between children in the number of hours spent in their pre-K (SD = 13) and K (SD = 8) classrooms.

Table 1
Descriptive statistics (N = 1015).

	% missing	n	M (SD)/%	Range
Child characteristics				
Race (White)	0%	424	42%	–
Race (Black)	0%	240	24%	–
Race (Latino)	0%	250	25%	–
Race (other)	0%	101	10%	–
Gender (male)	0%	496	49%	–
Family characteristics				
Poverty status	6%	952	57%	–
Maternal education (years)	2%	996	12.52 (2.20)	8–20
Pre-K classroom characteristics				
Hours per week	4%	974	24.19 (13.01)	6.5–60
CLASS IS	0%	1015	2.20 (0.79)	1–4.47
CLASS ES	0%	1015	5.26 (0.72)	2.92–6.87
CLASS CO	0%	1015	4.42 (0.82)	2.13–6.54
K classroom characteristics				
Hours per week	15%	867	27.42 (8.32)	7.5–57.5
CLASS IS	13%	885	1.98 (0.67)	1–5.21
CLASS ES	13%	885	5.45 (0.70)	2.46–6.89
CLASS CO	13%	885	4.67 (0.68)	1.58–6.53
Child outcomes: language				
PPVT-3, fall of pre-K	21%	805	92.31 (14.02)	32–134
PPVT-3, spring of K	17%	838	97.35 (12.69)	54–132
OWLS Oral Expression Scale, fall of pre-K	22%	789	91.08 (12.24)	52–131
OWLS Oral Expression Scale, spring of K	16%	857	94.69 (12.93)	61–135
Child outcomes: literacy				
WJ III Letter–Word Identification, spring of K	15%	859	106.47 (12.28)	63–153
CTOPP Phoneme Elisions, spring of K	30%	710	4.67 (3.06)	0–19
Child outcomes: mathematics				
WJ III Applied Problems, fall of pre-K	22%	790	97.14 (13.14)	51–135
WJ III Applied Problems, spring of K	15%	859	101.31 (11.06)	53–144

Note: K = Kindergarten. CLASS = Classroom Assessment Scoring System. IS = Instructional Support. ES = Emotional Support. CO = Classroom Organization. PPVT-3 = Peabody Picture Vocabulary Test, 3rd edition. WJ = Woodcock–Johnson. CTOPP = Comprehensive Test of Phonological Processing. Standard scores (M = 100, SD = 15) are reported for the PPVT-3, OWLS, WJ III Letter–Word ID, and WJ III Applied Problems measures. Raw scores are reported for the CTOPP measure.

3.2. Measures

3.2.1. Classroom quality

The Classroom Assessment Scoring System (CLASS; Pianta et al., 2008) was used to assess classroom quality in pre-K and K. Trained data collectors conducted live observations in 30-minute cycles on multiple days in children's pre-K and K classrooms. CLASS dimension scores from all observation days within each school year were averaged in the current analyses to create dimension scores representing pre-K and K classroom quality. The CLASS dimensions were measured on a 1–7 scale, with scores from 1 to 2 representing low-quality, 3–5 representing mid-quality, and 6–7 representing high-quality.

An early version of the CLASS was used for the NCEDL data collection. For the current study, nine CLASS dimensions were aggregated into the three domain scores. The *Instructional Support* domain score is a composite of two dimensions (Concept Development and Quality of Feedback), the *Emotional Support* domain score is a composite of three dimensions (Positive Classroom Climate, Negative Classroom Climate [reversed], and Teacher Sensitivity), and the *Classroom Organization* domain score is a composite of three dimensions (Behavior Management, Productivity, and Instructional Learning Formats). Detailed descriptions of each CLASS dimension are available from Pianta et al. (2008). The recent version of the CLASS measure contains two dimensions that did not exist when the NCEDL data were collected, including Language Modeling and Regard for Student Perspectives, which were added to the Instructional Support and Emotional Support domains, respectively. However, the domain scores used in the current study are directly comparable to those used by Downer et al. (2012) in a previous study using the NCEDL data. Additionally, after the NCEDL data were collected, a version of the CLASS was developed for K through third grade classrooms, but was not used in the current study.

The standardized Cronbach's alpha coefficients for the domain scores in this sample during pre-K and K indicate adequate reliability for Instructional Support (α = 0.89 and 0.86), Emotional Support (α = 0.88 and 0.89), and Classroom Organization (α = 0.80 and 0.89). Pairwise correlations among the CLASS domain scores are provided in Table 2. The Instructional Support domain score was moderately correlated with Emotional Support and Classroom Organization domain scores during pre-K (r = 0.51–0.60) and K (r = 0.30–0.44), while the Emotional Support and Classroom Organization domain

Table 2
Correlations between classroom quality variables.

	(1)	(2)	(3)	(4)	(5)	(6)
Pre-K classroom quality						
1. CLASS IS	1.00					
2. CLASS ES	0.51	1.00				
3. CLASS CO	0.60	0.83	1.00			
K classroom quality						
4. CLASS IS	0.11	0.05	0.01	1.00		
5. CLASS ES	0.19	0.21	0.18	0.30	1.00	
6. CLASS CO	0.12	0.11	0.13	0.44	0.74	1.00

Note: All correlations at or above 0.11 are statistically significant below the $p < 0.05$ level. K = Kindergarten. CLASS = Classroom Assessment Scoring System. IS = Instructional Support. ES = Emotional Support. CO = Classroom Organization.

Table 3
Correlations between language, literacy, and mathematics variables.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Language								
1. PPVT-3, fall of pre-K	1.00							
2. PPVT-3, spring of K	0.68	1.00						
3. OWLS Oral Expression Scale, fall of pre-K	0.62	0.57	1.00					
4. OWLS Oral Expression Scale, spring of K	0.58	0.62	0.61	1.00				
Literacy								
5. WJ III Letter–Word Identification, spring of K	0.23	0.26	0.22	0.36	1.00			
6. CTOPP Phoneme Elisions, spring of K	0.29	0.33	0.29	0.43	0.47	1.00		
Mathematics								
7. WJ III Applied Problems, fall of pre-K	0.55	0.49	0.56	0.59	0.33	0.36	1.00	
8. WJ III Applied Problems, spring of K	0.37	0.44	0.38	0.49	0.43	0.42	0.53	1.00

Note: All correlations are statistically significant below the $p < 0.05$ level. K = Kindergarten. PPVT-3 = Peabody Picture Vocabulary Test, 3rd edition. WJ = Woodcock–Johnson.

scores were correlated at an even greater rate during pre-K ($r = 0.83$) and K ($r = 0.74$).

The classrooms in this sample generally scored in the low-quality range on the Instructional Support domain during pre-K ($M = 2.20$, $SD = 0.79$) and K ($M = 1.98$, $SD = 0.67$); in the mid- to high-quality range on the Emotional Support domain during pre-K ($M = 5.26$, $SD = 0.72$) and K ($M = 5.45$, $SD = 0.70$); and in the mid- to high-quality range on the Classroom Organization domain during pre-K ($M = 4.42$, $SD = 0.82$) and K ($M = 4.67$, $SD = 0.68$). Scores in the low-quality range for the Instructional Support domain are commonly observed in pre-K samples (Choi et al., 2016; Hamre et al., 2014; Moiduddin et al., 2012), while mid- to high-quality scores of 5, 6, or 7 are rarely observed (Pianta, Downer, & Hamre, 2016). However, Pianta et al. (2016) have found that scores of 3 and 4 on the Instructional Support scale are consistently associated with higher levels of student achievement. It is worthwhile to note that 17% of pre-K and 7% of K classrooms in our sample reported scores at or above 3 on the Instructional Support domain.

3.2.2. *Child outcomes*

Children's academic skills were assessed during the fall of pre-K and again in the spring of K using standardized assessment measures. Consistent with prior research using the NCEDL data (Burchinal et al., 2008), composite measures of children's language and literacy, but not mathematics skills were created. Internal-consistency coefficients for the individual measures were examined at each assessment period and reported in the NCEDL study description and documentation materials (Clifford, Bryant et al., 2005). Pairwise correlations among the child outcome measures are provided in Table 3.

3.2.2.1. Language. The Peabody Picture Vocabulary Test, 3rd edition (PPVT-3; Dunn & Dunn, 1997) is a standardized test of receptive vocabulary. During this test, children were shown a series of pictures on a page and asked to point to the picture that best represented the meaning of a word spoken by the assessor. Reliability coefficients were $\alpha = 0.82$ and $\alpha = 0.84$ during the pre-K and K assessment periods, respectively (Clifford, Bryant et al., 2005). The Oral Expression Scale from the Oral and Written Language Scale (OWLS; Carrow-Woolfolk, 1995) is a standardized test of children's understanding and use of expressive language. During this test, children were asked to look at a card containing one or more pictures while they were read a verbal stimulus out loud. Each stimulus prompted children to answer a question, complete a sentence, or create a new sentence or sentences. Reliability coefficients were $\alpha = 0.76$ and $\alpha = 0.73$ during the pre-K and K assessment periods, respectively (Clifford, Bryant et al., 2005). Standard scores ($M = 100$, $SD = 15$) were used in the analyses for both the PPVT-3 and OWLS. A composite measure of children's Language skills at the fall of pre-K and spring of K was calculated as the average of Z-scores ($M = 0$, $SD = 1$) from the PPVT-3 and OWLS during each assessment period.

3.2.2.2. Literacy. The Letter–Word Identification (ID) subtest of the Woodcock–Johnson (WJ) III (Woodcock, McGrew, & Mather, 2001) is a standardized measure of children's ability to identify letters and correctly pronounce written words. The measure developers report reliability coefficients ranging between $\alpha = 0.92$ and 0.94 for the 3- to 6-year-old age group. The Phoneme Elisions Subtest of the Comprehensive Test of Phonological Processing (CTOPP; Wagner, Torgesen, & Rashotte, 1999) is a standardized test that was used to assess children's phonemic awareness (i.e., the ability to detect individual sounds in words). During this test, children are asked to separate compound words and remove particular sounds. The measure developers report a test-retest reliability of $r = 0.88$ for children in the 5- to 7-year age range. In the analyses, standard scores ($M = 100$, $SD = 15$) were used for the WJ III Letter–Word ID and raw scores were used for the CTOPP with a possible range of 0–19. A spring of K Literacy composite was calculated as the average of Z-scores on the WJ III Letter–Word ID and the CTOPP. A fall of pre-K Literacy composite could not be calculated because the WJ III Letter–Word ID measure was not administered during this assessment period and the CTOPP was too difficult for most 4-year-olds to establish a baseline at this time point.

3.2.2.3. Mathematics. The Applied Problems subtest of the WJ III (Woodcock et al., 2001) was used to assess children's mathematical reasoning in the fall of pre-K and spring of K. During this test, children were asked to identify quantities and solve relatively simple mathematical word problems of increasing difficulty. Reliability coefficients for the scores on this measure were $\alpha = 0.76$ at the pre-K and K assessment periods (Clifford, Bryant et al., 2005). Additionally, the measure developers reported a one-year test-retest reliability of $r = 0.92$ for children in the 4- to 7-year-old age range (McGrew & Woodcock, 2001). Standard scores ($M = 100$, $SD = 15$) were used for the WJ III Applied Problems subtest in the analyses.

3.2.3. *Covariates*

Several demographic and classroom covariates were included in the analyses. Demographic covariates included gender (1 = male, 0 = female), child race (binary variables coded as White, African American, Latino, and other; with 1 = yes, 0 = no), family poverty status, which was defined as below 150% of the federal poverty line (1 = poor, 0 = not poor), and a numeric variable indicating mother's years of education. Classroom covariates included the state in which the pre-K program was located as well as the number of hours per week of operation for children's pre-K and K programs.

Additionally, we attempted to reduce potential selection bias in our analyses by including children's fall pre-K scores as a covariate in analyses predicting children's spring of K scores. Children's fall pre-K scores on the Language composite were used as the pre-K score covariate in models predicting spring of K Language and Literacy composite scores. For mathematics, children's fall pre-K scores on the Applied Problems measure were used as the pre-K score covariate in models predicting spring of K Applied Problems scores.

3.3. Data analyses

3.3.1. Missing data

Data were missing across several variables, including family poverty status (6%), maternal education (2%), hours per week in pre-K (4%) and K (15%), and each of the K CLASS variables (13%). Data were also missing for the child outcome measures, which ranged from 15% to 30% across all measures. There were no data missing on the gender, race, pre-K CLASS, or state variables. Descriptive statistics and the proportion of missingness on each analytic variable are included in Table 1. In the analyses, multiple imputation (MI) methods were used to retain cases with missing values on the analytic variables. In comparison to listwise deletion, MI has been shown to provide more valid estimates of associations observed in the data and minimize the potential for biased results when the reason data are missing is not a function of the outcome itself (i.e., ignorably missing; Allison, 2002). The MI procedure in SAS® (version 9.3) was used to create 20 datasets with imputed values for missing cases. Each of the individual child outcome measures were included in the MI model and the composite Language and Literacy measures were calculated based on the imputed datasets. All other analytic variables employed in this study were also included in the MI model along with several auxiliary variables. Statistical analyses were then conducted in SAS on each imputed dataset and the MIANALYZE procedure was used to compute univariate parameter estimates and associated standard errors from values generated with each imputed dataset.

3.3.2. Analyses

The present study used a multi-level regression framework to examine child and pre-K classroom level predictors of children's academic outcomes in the spring of K. Separate models for the spring of K language, literacy, and mathematics outcomes were examined using the MIXED procedure in SAS. In these two-level models, children (Level 1; i) were nested within their pre-K classrooms (Level 2; j) to account for the cluster sampling study design, which randomly sampled as many as four children per pre-K classroom. Before the analyses were conducted, we examined the nesting of study children in K classrooms. We found that 85% of the K classrooms contained only one study child, 12% contained two study children, and only 3% contained three or more study children. We also examined three-level models with the addition of a random effect term for K classrooms. However, the random effect was equal to zero, which indicated that there was no variance in the child outcome measures associated with the nesting of children in K classrooms. Therefore, we chose not to consider this level of nesting in our subsequent analyses.

The following generalized equation represents the multi-level model used to test each child outcome:

$$
\begin{aligned}
\Upsilon_{ij} &= \beta_0 + \beta_1\,\text{pre-K score}_{ij} + \beta_2\,\text{CLASS pre-K}_j + \beta_3\,\text{CLASS K}_{ij} \\
&+ \beta_4\,\text{CLASS pre-K} \times \text{CLASS K}_{ij} + \beta_{5-11}\,\text{covariates}_{ij} \\
&+ \beta_{12-17}\,\text{covariates}_j + \varepsilon_{ij} + \zeta_j
\end{aligned}
$$

In this generalized equation, Υ_{ij} is an academic outcome (language, literacy, or mathematics) for child i in pre-K classroom j; β_0 is the random between-child intercept; β_1 is the effect of the pre-K score covariate (i.e., child scores on the pre-K Language composite or Applied Problems) that varies at level 1; β_2 is the main effect of a pre-K CLASS domain score (Instructional Support, Emotional Support, or Classroom Organization) that varies at level 2; β_3 is the main effect of a K CLASS domain score (Instructional Support, Emotional Support, or Classroom Organization) that varies at level 1; β_4 is the effect of the cross-level interaction between a pre-K and K CLASS domain score used to test for the moderation of pre-K classroom quality by K classroom quality, which varies at level 1; $\beta_{5–11}$ are the effects of covariates that vary at level 1 (i.e., gender, three dummy coded race variables indicating child race, poverty status, maternal education, K hours per week); $\beta_{12–17}$ are the effects of covariates that vary at level 2 (i.e., pre-K hours per week and five dummy coded variables to indicate each state, with the sixth and largest state as the reference category)[1]; ε_{ij} is the level 1 residual term; and ζ_j is the level 2 random effect. All of the dependent variables and continuous independent variables were standardized to have a mean of 0 and standard deviation of 1 in order to report standardized regression coefficients in Table 4, which can be interpreted as effect sizes (Hedges, 2008), while the binary predictor variables were centered, but not standardized (i.e., gender, child race, poverty status, state). Standardizing and centering also enables us to provide comparable main effect coefficients to models without interactions. Therefore, the main effect coefficients for the pre-K and K CLASS domain scores (β_2 and β_3) reported in Table 4 did not differ substantially in magnitude or statistical significance from models that excluded the cross-level interaction between the pre-K and K CLASS domain scores (β_4).[2] Three dummy coded variables were included in the models to indicate child race, including Black (1 = yes), Latino (1 = yes), and Other (1 = yes), with White (1 = yes) as the reference category. An omnibus F-test was used to evaluate if the coefficients for Black, Latino, and Other jointly contributed to each model, as reported in Table 4.

For the analyses, we chose to examine separate models for each CLASS domain score (Instructional Support, Emotional Support, and Classroom Organization) in relation to each child outcome of interest (language, literacy, and mathematics). This approach was consistent with prior studies using the NCEDL data (Burchinal et al., 2008, 2010) and allowed us to assess the predictive power of each CLASS domain without interference from the other domains. Issues associated with multicollinearity can make it impossible to disentangle the predictive power of each CLASS domain when these highly-correlated variables are simultaneously examined in the same analytic model (Hamre et al., 2014). This approach does not allow us, however, to compare the relative predictive power between domains. Models 1, 2, and 3 presented in Table 4 included pre-K and K domain scores for the CLASS Instructional Support, Emotional Support, and Classroom Organization domains, respec-

[1] Coefficients associated with the state identification variable were included in all models, but are not presented in the results section or tables because reporting state-level statistics of any type is prohibited by the data use agreement for this dataset.

[2] Main effect coefficients for the pre-K and K CLASS domain scores from models that excluded the cross-level interaction can be made available from the corresponding author upon request.

Table 4
Multi-level regression model results predicting spring of kindergarten outcomes.

	Language		Literacy		Mathematics	
Parameter	β	(*SE*)	β	(*SE*)	β	(*SE*)
Model 1						
Covariates						
Race	$F = 1.24$		$F = 5.49^{***}$		$F = 1.10$	
Male	0.02	(0.04)	0.00	(0.05)	−0.01	(0.06)
Poverty status	−0.08	(0.05)	−0.07	(0.06)	−0.02	(0.07)
Maternal education (years)	0.06^{**}	(0.02)	0.06	(0.03)	0.05	(0.04)
Hours per week, pre-K	0.03	(0.02)	0.01	(0.03)	0.02	(0.03)
Hours per week, K	−0.03	(0.02)	0.06	(0.03)	0.09^{*}	(0.03)
Language, fall of pre-K	0.77^{***}	(0.03)	0.41^{***}	(0.04)	–	–
Mathematics, fall of pre-K	–	–	–	–	0.55^{***}	(0.03)
CLASS Instructional Support (IS)						
CLASS IS Pre-K	0.04^{*}	(0.02)	0.06^{*}	(0.03)	0.04	(0.03)
CLASS IS K	0.08^{**}	(0.02)	0.12^{***}	(0.03)	0.07^{*}	(0.04)
CLASS IS Pre-K × CLASS IS K	0.00	(0.02)	0.04	(0.03)	0.06^{*}	(0.03)
Model 2						
CLASS Emotional Support (ES)						
CLASS ES Pre-K	0.03	(0.02)	0.08^{*}	(0.03)	0.05	(0.03)
CLASS ES K	0.04	(0.02)	0.08^{*}	(0.03)	−0.02	(0.03)
CLASS ES Pre-K × CLASS ES K	0.03	(0.02)	0.03	(0.03)	0.06^{*}	(0.03)
Model 3						
CLASS Classroom Organization (CO)						
CLASS CO Pre-K	0.03	(0.02)	0.08^{*}	(0.03)	0.06	(0.03)
CLASS CO K	0.05^{*}	(0.02)	0.10^{**}	(0.03)	0.02	(0.03)
CLASS CO Pre-K × CLASS CO K	0.00	(0.02)	0.02	(0.03)	0.02	(0.03)

Note: K = Kindergarten. CLASS = Classroom Assessment Scoring System.
* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

tively. Coefficients for the covariates presented in Table 4 were derived from Model 1.[3]

4. Results

To test for the additive cumulative effect of classroom quality, we examined the main effects of pre-K and K classroom quality within each CLASS domain (i.e., β_2 CLASS pre-$K_j + \beta_3$ CLASS K_{ij}). To test for the multiplicative cumulative effect of classroom quality, we examined the interaction term between pre-K and K classroom quality within each CLASS domain (i.e., β_4 CLASS pre-K × CLASS K_{ij}). The results of these analyses are outlined below and presented in Table 4.

4.1. Language

Children's language skills in the spring of K were positively associated with the main effects of CLASS Instructional Support during pre-K ($\beta = 0.04$, $p = 0.04$) and K ($\beta = 0.08$, $p = 0.001$), but not their interaction ($\beta = 0.00$, $p = 0.87$). Thus, Instructional Support quality in pre-K and K was associated with children's language skills in an additive manner. Children's language skills were also positively associated with the main effect of CLASS Classroom Organization in K ($\beta = 0.05$, $p = 0.03$), but not in pre-K ($\beta = 0.03$, $p = 0.18$), nor their interaction ($\beta = 0.00$, $p = 0.81$). Language skills were not related to the main effects of CLASS Emotional Support in pre-K ($\beta = 0.03$, $p = 0.18$) and K ($\beta = 0.04$, $p = 0.11$), nor their interaction ($\beta = 0.03$, $p = 0.09$).

4.2. Literacy

Children's literacy skills in the spring of K were positively associated with the main effects of CLASS Instructional Support during pre-K ($\beta = 0.06$, $p = 0.03$) and K ($\beta = 0.12$, $p < 0.001$), but not their interaction ($\beta = 0.04$, $p = 0.16$); with the main effects of CLASS Emotional Support in pre-K ($\beta = 0.08$, $p = 0.02$) and K ($\beta = 0.08$, $p = 0.02$), but not their interaction ($\beta = 0.03$, $p = 0.27$); and with the main effects of CLASS Classroom Organization in pre-K ($\beta = 0.08$, $p = 0.01$) and K ($\beta = 0.10$, $p = 0.001$), but not their interaction ($\beta = 0.02$, $p = 0.56$). Thus, each domain of CLASS quality in pre-K and K was associated with children's literacy skills in an additive manner.

4.3. Mathematics

Children's mathematics skills in the spring of K were positively associated with the main effect of CLASS Instructional Support in K ($\beta = 0.07$, $p = 0.048$), but not in pre-K ($\beta = 0.04$, $p = 0.25$). Additionally, children's mathematics skills were not associated with the main effects of CLASS Emotional Support in pre-K ($\beta = 0.05$, $p = 0.08$) or K ($\beta = -0.02$, $p = 0.46$). However, we did find statistically significant interactions between pre-K and K Instructional Support ($\beta = 0.06$, $p = 0.03$) as well as pre-K and K Emotional Support ($\beta = 0.06$, $p = 0.02$), which both indicated that the predictive effect of pre-K quality increased by 0.06 standard deviation units with every one standard deviation unit increase in K quality. In a sensitivity analysis, we probed the interaction between pre-K and K Instructional Support and found that the predictive effect of pre-K Instructional Support was statistically significant for children who also experienced *high* levels of Instructional Support quality in K (i.e., at +1 *SD*, $\beta = 0.10$, $SE = 0.04$, $p = 0.02$), but was not statistically significant for children who experienced *average* (i.e., at the mean; $\beta = 0.04$, $SE = 0.03$, $p = 0.25$) or *low* levels of Instructional Support quality in K (i.e., at −1 *SD*, $\beta = -0.03$, $SE = 0.04$, $p = 0.49$). Similarly,

[3] Coefficients for the covariates presented in Table 4 did not differ substantially from coefficients derived from the models including the Emotional Support or Classroom Organization domain scores, which can be made available from the corresponding author upon request.

after probing the interaction between pre-K and K Emotional Support, we found that the predictive effect of pre-K Emotional Support was statistically significant for children who also experienced high levels of Emotional Support in K (i.e., at +1 *SD*; $\beta = 0.11$, $SE = 0.04$, $p = 0.007$), but was not statistically significant for children who experienced average (i.e., at the mean; $\beta = 0.05$, $SE = 0.03$, $p = 0.08$) or low levels of Emotional Support quality in K (i.e., at −1 *SD*; $\beta = -0.01$, $SE = 0.04$, $p = 0.80$). Additionally, mathematics skills were not significantly related to the main effects of CLASS Classroom Organization in pre-K ($\beta = 0.06$, $p = 0.07$) and K ($\beta = 0.02$, $p = 0.58$), nor their interaction ($\beta = 0.02$, $p = 0.56$).

Thus, based on these interactions, we found evidence to support both propositions of dynamic complementarity in relation to children's mathematics skills, because the positive predictive effect of pre-K Instructional Support and Emotional Support was enhanced as the level of K classroom quality increased in these domains, but the positive predictive effect of pre-K classroom quality was only found to be statistically significant for children who subsequently experienced high levels of K quality. It is worth noting that we probed the moderating effect of K classroom quality at three conventional values (i.e., +1 *SD*, mean, and −1 *SD*). Yet, the continuous-by-continuous interactions used in our analyses reflect the full range of K classroom quality in our sample, indicating that even larger effects of pre-K classroom quality could be evident for children in K classrooms with quality scores greater than +1 *SD*.

5. Discussion

The current study advances research about early childhood education through an examination of the cumulative effects of classroom environments during pre-K and K in relation to children's language, literacy, and mathematics skills in the spring of K. We found that additive cumulative effects described associations between classroom quality in relation to children's language and literacy skills: higher-quality pre-K uniformly predicted better skills at all levels of K quality, but higher-quality K also contributed in an additive manner. These findings illustrate how children profited from higher-quality classroom environments during both pre-K and K. Alternatively, we found evidence of a multiplicative cumulative effect of classroom quality in relation to children's mathematics skills, such that the predictive effect of higher-quality pre-K classroom environments was enhanced as the level K classroom quality increased. Moreover, the positive predictive effect of pre-K classroom quality was only statistically significant for children who subsequently experienced *high*, but not *average* or *low* levels of K classroom quality. Our findings related to mathematics illustrate the two propositions of *dynamic complementarity* as described by Cunha and Heckman (2007): (1) later investments in children's skill development bolster the productivity of earlier investments and (2) earlier investments do not continue to be productive unless they are followed up by later investments.

Thus, the accumulation of higher-quality classroom experiences mattered in different ways for different academic skills. We posit that the multiplicative cumulative effects were observed in relation to mathematics skills, but not language and literacy skills because the links between pre-K classroom quality and children's language and literacy skills were initially more robust and, thus, less dependent on subsequent classroom quality. Indeed, prior research using the NCEDL dataset has documented larger effects of pre-K classroom quality on children's end of pre-K language and literacy skills ($d = 0.06$–0.11) compared to mathematics skills ($d = 0.05$–0.06; Howes et al., 2008). It is also possible that the development of mathematics skills requires more sustained support and may be more difficult for children to acquire without exposure to high-quality classroom environments across grades.

Previous studies have examined preschool classroom quality in relation to children's concurrent and long-term academic outcomes; documenting positive associations during the preschool period that begin to fade out during the early elementary grades. Additionally, a separate body of research has examined the quality of children's classroom experiences in elementary school—including some of our own work that has documented the additive cumulative effect of better-quality classroom environments between K and third grade in relation to children's third grade literacy skills (Vernon-Feagans et al., 2018). However, few studies have documented the accumulation of higher-quality classroom experiences during preschool and elementary school. Many scholars have argued that educational experiences in elementary school are crucial for sustaining the long-term impacts of preschool programming and therefore, the contributions of preschool and subsequent elementary school experiences should be jointly considered (Bailey, Duncan, Odgers, & Yu, 2016; Phillips et al., 2017).

The current study added important evidence to the extant research literature on the accumulation of higher-quality classroom experiences across the transition to school. Prior studies have documented the benefit of cumulative classroom emotional support/classroom organization quality during pre-K and K in relation to children's social–emotional skills (Broekhuizen et al., 2016) as well as the cumulative benefit of emotional support quality during child care and elementary school in relation to academic skills (Ansari & Pianta, 2018b). By focusing on the cumulative benefit of classroom quality during pre-K and K in relation to children's early academic skills, we found that higher-quality classroom experiences during both years was associated with higher levels of language, literacy, and mathematics skills for children, albeit the pattern of these associations was additive for language and literacy versus multiplicative for mathematics.

Additionally, our findings related to mathematics illustrate the importance of high-quality classroom environments in K as a potential mechanism for sustaining the benefit of high-quality pre-K and also provides some support for the concept of dynamic complementarity, which is derived from Cunha and Heckman's (2007) cumulative model of human capital acquisition. When applied to research on early childhood education, this model has received some support from earlier studies of preschool classroom quality (Ansari & Pianta, 2018b; Broekhuizen et al., 2016) as well as preschool program participation (Ansari & Pianta, 2018a; Bierman et al., 2014; Jenkins et al., 2018; Zhai et al., 2012). Given the emerging empirical support for this model in the field of early childhood education, it is possible that the concept of dynamic complementarity can and should be applied to a wider range of educational and developmental research areas.

5.1. Implications for future research and practice

Education stakeholders are increasingly interested in designing policies and practices to align classroom instruction across pre-K through third grade in order to promote continuity in children's educational experiences (Bogard & Takanishi, 2005; Phillips et al., 2017; Pianta et al., 2016; Ritchie & Gutmann, 2014; Stipek, Clements, Coburn, Franke, & Farran, 2017). Several studies provide evidence linking instructional alignment to children's developmental outcomes (Jenkins et al., 2018; Mattera, Jacob, & Morris, 2018; Zellman & Kilburn, 2015). To date, most instructional alignment initiatives have focused on the use of curricula, instructional practices, learning standards, and assessments. Although the current study did not examine these important facets of instructional alignment, Stipek et al. (2017) have acknowledged that continuity in instruction depends on the quality of instruction. The current study illustrates how the accumulation of high-quality classroom experiences across pre-K and K can promote children's academic learning gains during this period. Future research into instructional alignment should also consider the underlying role of high-quality teacher–child interactions, cumulative or not. For example, does instructional alignment provide greater benefits for children's learning in higher- compared to lower-quality classroom contexts?

Given the wealth of evidence to support the developmental importance of high-quality teacher–child interactions, instructional alignment initiatives could also include strategies to promote the quality of teacher–child interactions (e.g., professional development programs) in addition to strategies promoting traditional forms of instructional alignment (e.g., curricula, instructional practices, standards, and assessments). Moreover, strategies to promote teacher–child interactions might focus on aspects of quality related to Instructional Support, given that this domain of quality was rated the lowest in pre-K and K, but was found to be the most consistent predictor of child outcomes in our study.

Future research related to instructional alignment may also benefit from the development of new strategies and methods to assess the quality of teacher–child interactions across grades as these interactions naturally change and adapt to match children's developmental progressions through school. Unlike the concept of cumulative classroom quality that was considered in the current study, Stipek et al. (2017) describe how the term *continuity* implies that "instruction needs to change as children develop new understandings and skills, are able to take on new forms of participation, develop language, and take on new content" (p. 4). Relatedly, the type of teacher–child interactions that constitute high-quality instruction in pre-K may differ from the type of interactions that constitute high-quality instruction in K, first, second, or third grade and, if so, then quality measurement should capture the type of interactions that are most salient during each period. Although current study could not address the issue of *continuity* in classroom quality, we provide evidence to indicate that exposure to the same high-quality classroom processes in both pre-K and K—as measured by the CLASS—can benefit children's learning.

5.2. Strengths and limitations

Despite the important contributions of the current study, the results should be viewed in light of several limitations. First, no causal inferences can be made from this study because the findings are correlational in nature. To account for potential selection bias of higher-skilled children into higher-quality classrooms, we included child, demographic, and classroom covariates in the analytic models—including children's pre-K entry-level skills. However, the potential for selection bias was not completely eliminated in this study and the degree to which classroom quality matters for child outcomes may be attributable to unmeasured, non-random sorting of children into higher-quality classrooms in pre-K and K. Second, it is possible that important aspects of classroom process quality have gone unmeasured by the classroom observation instrument used in the current study. The early version of the CLASS measure used in this study did not include two dimension scores that have since been added to the most recent version of this measure (i.e., Language Modeling and Regard for Student Perspectives). Additionally, the current study did not examine other domain-specific measures designed to document classroom processes most proximal to children's skill development in the domains of language, literacy, and mathematics. Domain-specific measures could provide a more refined understanding of the type of classroom processes that boost children's learning across pre-K and elementary school. Third, although the current study aimed to examine the accumulation of *higher-quality* classroom environments, the average level of CLASS Instructional Support scores observed in this sample were in the low-quality range in pre-K ($M = 2.20$, $SD = 0.79$) and K ($M = 1.98$, $SD = 0.67$) as defined by the measure developers. However, scores in the low-quality range for Instructional Support are commonly observed in pre-K samples (Choi et al., 2016; Hamre et al., 2014; Moiduddin et al., 2012), while mid- to high-quality scores are rarely observed (Pianta et al., 2016). Nonetheless, the current study provides some evidence to suggest that exposure to moderate levels of Instructional Support quality during pre-K and K can boost children's learning gains across the transition to school. Fourth, although our study focused on associations between pre-K and K classroom quality in relation to child outcomes at the end of K, other analytic specifications could have been used to consider quality effects on child outcomes assessed at the end of pre-K and/or beginning of kindergarten using the NCEDL data. Finally, it is important to note that the NCEDL data used in the current study are now more than a decade old. However, these findings remain relevant to early childhood education programming today given that the magnitude of classroom quality effects found in the current study were similar to those found in more recent samples of children and classrooms (e.g., Weiland et al., 2013).

Despite these limitations, the current study has several strengths that should be noted. First, we used a large sample of socioeconomically diverse children recruited from pre-K programs across six states, which enhances the generalizability of these results. Second, we examined classroom quality in pre-K and K based on direct observations during each year. Third, we used multiple imputation methods to retain all cases in our analyses that were missing data on some analytic variables. Finally, the current study contributes new evidence to the growing research base on the importance of multiple years of high-quality classroom environments during early childhood.

6. Conclusion

The current study examined three domains of classroom quality in pre-K and K—Instructional Support, Emotional Support, and Classroom Organization—in relation to children's language, literacy, and mathematics skills in the spring of K. Our study controlled for a number of important factors that allowed us to estimate—with some confidence and precision—associations between classroom quality and children's academic skills. However, due to the observational nature of our study, no casual conclusions can be made.

The results of our study suggest that an accumulation of higher-quality classroom experiences during pre-K and K may be salient for the development of children's language, literacy, and mathematics skills across the transition to school. In relation to children's K language and literacy skills, we found that the predictive effect of K classroom quality added to the predictive effect of pre-K classroom quality. In relation to mathematics, we found that the predictive effect pre-K classroom quality was enhanced by K classroom quality, but was only statistically significant for children who experienced high levels of classroom quality in K, not average or low levels of K classroom quality. These findings contribute to a growing recognition of the importance of high-quality educational environments across the transition to school. We believe that future research should explore new strategies and methods to assess the quality of teacher–child interactions across grades as these interactions change and adapt to match children's developmental progressions through school. Additionally, we hope that future research will consider if high-quality teacher–child interactions (cumulative or not) enhance the implementation and efficacy of instructional alignment initiatives across the transition from pre-K to elementary school.

Acknowledgements

This research study was supported by a predoctoral fellowship provided by the National Institute of Child Health and Human Development (T32-HD07376) through the Center for Developmental Science, University of North Carolina at Chapel Hill, to the first author (R.C. Carr).

References

Allison, P. D. (2002). *Missing data*. Thousand Oaks, CA: SAGE Publications.
Anderson, S., & Phillips, D. (2017). Is pre-K classroom quality associated with kindergarten and middle-school academic skills? *Developmental Psychology*, *53*(6), 1063–1078. http://dx.doi.org/10.1037/dev0000312
Ansari, A., & Pianta, R. C. (2018a). The role of elementary school quality in the persistence of preschool effects. *Children and Youth Services Review*, *86*(1), 120–127. http://dx.doi.org/10.1016/j.childyouth.2018.01.025
Ansari, A., & Pianta, R. C. (2018b). Variation in the long-term benefits of child care: The role of classroom quality in elementary school. *Developmental Psychology*, *54*(10), 1–14. http://dx.doi.org/10.1037/dev0000513
Auger, A., Farkas, G., Burchinal, M. R., Duncan, G. J., & Vandell, D. L. (2014). Preschool center care quality effects on academic achievement: An instrumental variables analysis. *Developmental Psychology*, *50*(12), 2559–2571. http://dx.doi.org/10.1037/a0037995
Bailey, D., Duncan, G. J., Odgers, C. L., & Yu, W. (2016). Persistence and fadeout in the impacts of child and adolescent interventions. *Journal of Research on Educational Effectiveness*, *10*(1), 7–39. http://dx.doi.org/10.1080/19345747.2016.1232459
Bassok, D., Gibbs, C. R., & Latham, S. (2018). Preschool and children's outcomes in elementary school: Have patterns changed nationwide between 1998 and 2010? *Child Development*, *00*(0), 1–23. http://dx.doi.org/10.1111/cdev.13067
Belsky, J., Vandell, D. L., Burchinal, M. R., Clarke-Stewart, K. A., McCartney, K., & Owen, M. T. (2007). Are there long-term effects of early child care? *Child Development*, *78*(2), 681–701. http://dx.doi.org/10.2307/4139252
Bierman, K. L., Nix, R. L., Heinrichs, B. S., Domitrovich, C. E., Gest, S. D., Welsh, J. A., & Gill, S. (2014). Effects of Head Start REDI on children's outcomes 1 year later in different kindergarten contexts. *Child Development*, *85*(1), 140–159. http://dx.doi.org/10.1111/cdev.12117
Bogard, K., & Takanishi, R. (2005). PK-3: An aligned and coordinated approach to education for children 3 to 8 years old. *Social Policy Report*, *19*(3), 1–24. Retrieved from. https://www.fcd-us.org/pk-3-an-aligned-and-coordinated-approach-to-education-for-children-3-to-8-years-old/
Broekhuizen, M. L., Mokrova, I. L., Burchinal, M. R., Garrett-Peters, P. T., & The FLP Key Investigators. (2016). Classroom quality at pre-kindergarten and kindergarten and children's social skills and behavior problems. *Early Childhood Research Quarterly*, *36*(1), 212–222. http://dx.doi.org/10.1016/j.ecresq.2016.01.005
Bronfenbrenner, U., & Morris, P. A. (2006). The bioecological model of human development. In W. Damon, & R. M. Lerner (Eds.), *Handbook of child psychology, Vol. 1: Theoretical models of human development* (6th ed., pp. 793–828). Hoboken, NJ: Wiley.
Burchinal, M. R., Howes, C., Pianta, R. C., Bryant, D., Early, D., Clifford, R., & Barbarin, O. (2008). Predicting child outcomes at the end of kindergarten from the quality of pre-kindergarten teacher-child interactions and instruction. *Applied Developmental Science*, *12*(3), 140–153. http://dx.doi.org/10.1080/10888690802199418
Burchinal, M. R., Vandergrift, N., Pianta, R. C., & Mashburn, A. (2010). Threshold analysis of association between child care quality and child outcomes for low-income children in pre-kindergarten programs. *Early Childhood Research Quarterly*, *25*(2), 166–176. http://dx.doi.org/10.1016/j.ecresq.2009.10.004
Burchinal, M. R., Magnuson, K. A., Powell, D., & Soliday Hong, S. (2015). Early childcare and education. In R. M. Lerner, M. H. Bornstein, & T. Leventhal (Eds.), *Handbook of child psychology and developmental science* (7th ed., pp. 223–267). Hoboken, NJ: Wiley.
Burchinal, M. R., Xue, Y., Auger, A., Tien, H. C., Mashburn, A., Peisner-Feinberg, E. S., . . . & Tarullo, L. (2016). Testing for quality thresholds and features in early care and education. *Monographs of the Society for Research in Child Development*, *81*(2), 46–63. http://dx.doi.org/10.1111/mono.12238
Burchinal, M. R. (2018). Measuring early care and education quality. *Child Development Perspectives*, *12*(1), 3–9. http://dx.doi.org/10.1111/cdep.12260
Carrow-Woolfolk, E. (1995). *Oral and written language scales (OWLS)*. Circle Pines, MN: American Guidance Service.
Choi, J. Y., Elicker, J., Christ, S. L., & Dobbs-Oates, J. (2016). Predicting growth trajectories in early academic learning: Evidence from growth curve modeling with Head Start children. *Early Childhood Research Quarterly*, *36*(1), 244–258. http://dx.doi.org/10.1016/j.ecresq.2015.12.017
Clifford, R. M., Barbarin, O., Chang, F., Early, D., Bryant, D., Howes, C., . . . & Pianta, R. C. (2005). What is pre-kindergarten? Characteristics of public pre-kindergarten programs. *Applied Developmental Science*, *9*(3), 126–143. http://dx.doi.org/10.1207/s1532480xads0903_1
Clifford, R. M., Bryant, D., Burchinal, M., Barbarin, O., Early, D., Howes, C., . . . & Winton, P. (2005). *National center for early development and learning multistate study of pre-Kindergarten: Study description & documentation for public release of data*. Ann Arbor, MI: Child Care and Early Education Research Connections, Inter-University Consortium for Political and Social Research.
Cunha, F., & Heckman, J. (2007). The technology of skill formation. *The American Economic Review*, *97*(2), 31–47. http://dx.doi.org/10.1257/aer.97.2.31
Curenton, S. M., Dong, N., & Shen, X. (2015). Does aggregate school-wide achievement mediate fifth grade outcomes for former early childhood education participants? *Developmental Psychology*, *51*(7), 921–934. http://dx.doi.org/10.1037/a0039295
Downer, J. T., López, M. L., Grimm, K. J., Hamagami, A., Pianta, R. C., & Howes, C. (2012). Observations of teacher–child interactions in classrooms serving Latinos and dual language learners: Applicability of the Classroom Assessment Scoring System in diverse settings. *Early Childhood Research Quarterly*, *27*(1), 21–32. http://dx.doi.org/10.1016/j.ecresq.2011.07.005
Dunn, L. M., & Dunn, L. M. (1997). *Peabody Picture Vocabulary Test—Revised* (3rd ed.). Circle Pines, MN: American Guidance Services.
Early, D., Barbarin, O., Bryant, D., Burchinal, M. R., Chang, F., Clifford, R., . . . & Barnett, W. S. (2005). *Pre-kindergarten in eleven states: NCEDL's Multi-State Study of Pre-Kindergarten & Study of State-Wide Early Education Programs (SWEEP): Preliminary descriptive report.*. Retrieved from. http://fpg.unc.edu/node/4654
Hamre, B. K., Hatfield, B., Pianta, R. C., & Jamil, F. (2014). Evidence for general and domain-specific elements of teacher–child interactions: Associations with preschool children's development. *Child Development*, *85*(3), 1257–1274. http://dx.doi.org/10.1111/cdev.12184
Hedges, L. V. (2008). What are effect sizes and why do we need them? *Child Development Perspectives*, *2*(3), 167–171. http://dx.doi.org/10.1111/j.1750-8606.2008.00060.x
Hindman, A. H., Skibbe, L. E., Miller, A., & Zimmerman, M. (2010). Ecological contexts and early learning: Contributions of child, family, and classroom factors during Head Start, to literacy and mathematics growth through first grade. *Early Childhood Research Quarterly*, *25*(2), 235–250. http://dx.doi.org/10.1016/j.ecresq.2009.11.003
Howes, C., Burchinal, M. R., Pianta, R. C., Bryant, D., Early, D., Clifford, R., & Barbarin, O. (2008). Ready to learn? Children's pre-academic achievement in pre-kindergarten programs. *Early Childhood Research Quarterly*, *23*(1), 27–50. http://dx.doi.org/10.1016/j.ecresq.2007.05.002
Jenkins, J. M., Watts, T. W., Magnuson, K., Gershoff, E., Clements, D., Sarama, J., & Duncan, G. J. (2018). Do high quality kindergarten and first grade classrooms mitigate preschool fadeout? *Journal of Research on Educational Effectiveness*, *11*(3), 339–374. http://dx.doi.org/10.1080/19345747.2018.1441347
Keys, T. D., Farkas, G., Burchinal, M. R., Duncan, G. J., Vandell, D. L., Li, W., . . . & Howes, C. (2013). Preschool center quality and school readiness: Quality effects and variation by demographic and child characteristics. *Child Development*, *84*(4), 1171–1190. http://dx.doi.org/10.1111/cdev.12048
Magnuson, K. A., Ruhm, C., & Waldfogel, J. (2007). The persistence of preschool effects: Do subsequent classroom experiences matter? *Early Childhood Research Quarterly*, *22*(1), 18–38. http://dx.doi.org/10.1016/j.ecresq.2006.10.002
Mashburn, A. J., Pianta, R. C., Barbarin, O. A., Bryant, D., Hamre, B. K., Jason, T. D., . . . & Howes, C. (2008). Measures of classroom quality in prekindergarten and children's development of academic, language, and social skills. *Child Development*, *79*(3), 732–749. http://dx.doi.org/10.2307/27563514
Mattera, S. K., Jacob, R. T., & Morris, P. (2018). *Strengthening children's math skills with enhanced instruction: The impacts of making pre-k count and high 5s on kindergarten outcomes.*. Retrieved from. https://www.mdrc.org/publication/strengthening-children-s-math-skills-enhanced-instruction
McGrew, K. S., & Woodcock, R. W. (2001). *Woodcock–Johnson III technical manual*. Itasca, IL: The Riverside Publishing Company.
Moiduddin, E., Aikens, N., Tarullo, L., West, J., & Xue, Y. (2012). *Child outcomes and classroom quality in FACES 2009. OPRE Report 2012-37a*. Washington, DC: Office of Planning, Research and Evaluation (OPRE), Administration for Children and Families (ACF), U.S. Department of Health and Human Services.
NICHD Early Child Care Research Network. (2002). Early child care and children's development prior to school entry: Results from the NICHD Study of Early Child Care. *American Educational Research Journal*, *39*(1), 133–164. http://dx.doi.org/10.3102/00028312039001133
Peisner-Feinberg, E. S., Burchinal, M. R., Clifford, R. M., Culkin, M. L., Howes, C., Kagan, S. L., & Yazejian, N. (2001). The relation of preschool child-care quality to children's cognitive and social developmental trajectories through second grade. *Child Development*, *72*(5), 1534–1553. http://dx.doi.org/10.1111/1467-8624.00364
Phillips, D., Lipsey, M. W., Dodge, K. A., Haskins, R., Bassok, D., Burchinal, M. R., . . . & Weiland, C. (2017). *Puzzling it out: The current state of scientific knowledge on pre-kindergarten effects: A consensus statement*. Retrieved from. https://www.brookings.edu/research/puzzling-it-out-the-current-state-of-scientific-knowledge-on-pre-kindergarten-effects/
Pianta, R. C., La Paro, K. M., & Hamre, B. K. (2008). *Classroom Assessment Scoring System (CLASS)*. Baltimore, MD: Paul H. Brookes Publishing Co.
Pianta, R. C., Downer, J., & Hamre, B. K. (2016). Quality in early education classrooms: Definitions, gaps, and systems. In S. McLanahan, J. M. Currie, R. Haskins, C. E. Rouse, & I. Sawhill (Eds.), *Starting Early: Education from prekindergarten to third grade* (Vol. 26, 2) (pp. 119–137). The Future of Children.
Ritchie, S., & Gutmann, L. (Eds.). (2014). *FirstSchool: Transforming preK–3rd grade for African American, Latino, and low-income children*. New York, NY: Teachers College Press.

Stipek, D., Clements, D., Coburn, C., Franke, M., & Farran, D. (2017). PK-3: What does it mean for instruction? *SRCD Social Policy Report*, *30*(2), 1–23. https://www.srcd.org/sites/default/files/documents/spr_30_2_final.pdf

Vandell, D. L., & Wolfe, B. (2000). *Child care quality: Does it matter and does it need to be improved? (Report No. 78)*.. Retrieved from Institute for Research on Poverty, University of Wisconsin-Madison. https://aspe.hhs.gov/report/child-care-quality-does-it-matter-and-does-it-need-be-improved-full-report

Vandell, D. L., Burchinal, M., & Pierce, K. M. (2016). Early child care and adolescent functioning at the end of high school: Results from the NICHD Study of Early Child Care and Youth Development. *Developmental Psychology*, *52*(10), 1634–1645. http://dx.doi.org/10.1037/dev0000169

Vernon-Feagans, L., Mokrova, I. L., Carr, R. C., Garrett-Peters, P. T., Burchinal, M. R., & The FLP Key Investigators. (2018). Cumulative years of classroom quality from kindergarten to third grade: Prediction to children's third grade literacy skills. *Early Childhood Research Quarterly*, http://dx.doi.org/10.1016/j.ecresq.2018.06.005. Advanced online publication

Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1999). *Comprehensive test of phonological processing*. Austin, TX: PRO-ED.

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. http://dx.doi.org/10.1016/j.ecresq.2012.12.002

Wen, X., Bulotsky-Shearer, R. J., Hahs-Vaughn, D. L., & Korfmacher, J. (2012). Head Start program quality: Examination of classroom quality and parent involvement in predicting children's vocabulary, literacy, and mathematics achievement trajectories. *Early Childhood Research Quarterly*, *27*(4), 640–653. http://dx.doi.org/10.1016/j.ecresq.2012.01.004

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock–Johnson III tests of achievement*. Itasca, IL: The Riverside Publishing Company.

Yoshikawa, H., Weiland, C., Brooks-Gunn, J., Burchinal, M. R., Espinosa, L. M., Gromley, W. T., . . . & Zaslow, M. J. (2013). *Investing in our future: The evidence base on preschool education*. Washington, DC: Society for Research in Child Development. Retrieved from. http://fcd-us.org/resources/evidence-base-preschool

Zellman, G., & Kilburn, M. (2015). *How well did P–3 work? The Hawai'i preschool-third grade education reform initiative*. Santa Monica, CA: Rand Corporation. Retrieved from. www.rand.org/pubs/research_reports/RR1100.html

Zhai, F., Raver, C. C., & Jones, S. M. (2012). Academic performance of subsequent schools and impacts of early interventions: Evidence from a randomized controlled trial in Head Start settings. *Children and Youth Services Review*, *34*(5), 946–954. http://dx.doi.org/10.1016/j.childyouth.2012.01.026

Zucker, T. A., Cabell, S. Q., Justice, L. M., Pentimonti, J. M., & Kaderavek, J. N. (2013). The role of frequent, interactive prekindergarten shared reading in the longitudinal development of language and literacy skills. *Developmental Psychology*, *49*(8), 1425–1439. http://dx.doi.org/10.1037/a0030347