Early Childhood Research Quarterly 64 (2023) 290–301

Contents lists available at ScienceDirect

Early Childhood Research Quarterly

journal homepage: www.elsevier.com/locate/ecresq

Classroom skill compositions and preschoolers' early academic and executive function outcomes ☆

Qingqing Yang[a], Arya Ansari[a,b], Kelly M. Purtell[a,b], Robert C. Pianta[c], Jessica V. Whittaker[c], Virginia E. Vitiello[c]

[a] Department of Human Sciences, The Ohio State University
[b] Crane Center for Early Childhood Research and Policy, The Ohio State University
[c] Center for Advanced Study of Teaching and Learning, University of Virginia

ARTICLE INFO

Keywords:
preschoolers
peer effects
classmates' skills
within-domain
cross-domain

ABSTRACT

This study examined the links between classroom skill compositions and preschoolers' early learning and development in the nationally representative Head Start Family and Child Experiences Survey 2014 (FACES, $n = 1{,}711$ children/207 classrooms) and public pre-K programs in a county in Virginia ($n = 1{,}467$ children/123 classrooms). Results from a series of covariate-adjusted multilevel regression models revealed that there were small within-domain associations between classmates' skill mean and individual children's academic and executive function skill development in FACES, but not in the Virginia data. There were no consistent associations across skill domains nor as a function of classmates skills' heterogeneity. In addition, we found little evidence that these associations between classroom skill compositions and individual children's development varied by children's initial skill levels, family income, maternal education, and home language. When taken together, these findings inform the discourse on peer influences on children's learning in early childhood.

Considerable research has shown that children's classmates are an important influence on their educational experiences (Henry & Rickman, 2007; Skibbe et al., 2012). These influences may begin to emerge in preschool when children spend a large amount of time with peers with different levels of academic and executive function (EF) skills in either group activities or free play (Early et al., 2010; Justice et al., 2021). Classmates' skills can vary in two ways, both with respect to the mean of each skill and the degree to which individual classmates' skills deviate from the classroom mean (i.e., heterogeneity; Atkins-Burnett et al., 2017; Finch et al., 2019). These two aspects of classmates' skills, referred to as classroom skill composition, have implications on children's skill gains either through peer-peer learning or bringing changes in the classroom learning environment (e.g., changing teachers' provision of instruction and tasks).

The importance of classmates' skills is supported by work that has considered classmates' average language skills as a predictor of preschoolers' learning outcomes (Justice et al., 2011; Mashburn et al., 2009). However, less attention has been paid to other domains of development, including children's literacy, mathematics, and executive function skills, which are key correlates of long-term school success (Duncan et al., 2007; Reynolds et al., 2011). Existing studies of the association between classrooom skill compositions and children's development has also generated mixed findings, with very few examining the role of skill heterogeneity in the preschool context (e.g., Justice et al., 2011; Mashburn et al., 2009). Evidence is also mixed when considering to what extent classroom skill composition matters differently for different subgroups of children who enter preschool with diverse experiences (Kohl et al., 2021). Thus, this study sought to address these gaps in knowledge by examining the extent to which each aspect of skill composition within preschool classrooms (i.e., mean and heterogeneity) are associated with children's gains in language, literacy, mathematics, and executive function. As part of this effort, we also: (a) test for both within- and cross-domain associations and (b) examine the extent to which these associations vary based on children's school-entry skills, maternal education, household income, and home language. To help resolve the mixed findings in prior studies, we leveraged two large and contemporary datasets collected between 2010-2020 in the United States and performed parallel analyses to answer the above questions.

☆ We gratefully acknowledge the support of our many partners: school district leaders, community programs, teachers, parents, and children. Their enthusiastic cooperation and participation made much of this work possible. We also extend appreciation and recognition to Marcia Kraft-Sayre, Marianna Lyulchenko, Laura Helferstay, and Brittany Kerr, who each made valuable contributions to the project. The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305N160021 to the University of Virginia. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

https://doi.org/10.1016/j.ecresq.2023.04.003
Received 29 June 2022; Received in revised form 31 March 2023; Accepted 8 April 2023
Available online 20 April 2023
0885-2006/

The Role of Peers in Child Development

The availability of preschool in the United States has expanded in recent years (National Center for Education Statistics, 2018), resulting in a greater number of 3- and 4-year-olds who spend a significant portion of their day in a classroom with peers. Although there is variability from classroom to classroom, preschoolers within the United States spend, on average, almost a third of the school day in free-choice play and a little over a third of the day in teacher-assigned activities (Early et al., 2010). Similar research also finds that on average, over half of their day is spent in large-group, small-group, or dyad activities (Justice et al., 2021). Within these activities, the time preschoolers spend talking with their peers (18%) is similar to the time they spend speaking with adults (17%; Dickinson & Tabors, 2001). Moreover, many teacher-assigned activities are organized in either a small- or large-group setting where children have many opportunities to interact with peers and observe peers' behaviors, and teachers adjust their involvement and instruction based on the group's performance (Booren et al., 2012; Fuligni et al., 2012). As such, children's peers are a key part of the classroom context and may contribute to their development.

Previous theories and studies have highlighted both direct and indirect pathways through which peers may influence children's learning and development. According to social cognitive theory and sociocultural theory, interacting with classmates provides children with the opportunity to observe and model classmates' behaviors and skills and to scaffold their classmates' learning, both of which facilitate and reinforce their own learning (Bandura, 1986; Vygotsky, 1978). For example, studies have shown that preschoolers learn early academic skills (e.g., naming sight words) taught to their more skilled peers by modeling their peers' skills in small-group activities (Ledford & Wolery, 2015). Peer effects may also manifest by shaping the learning environment within a classroom. For instance, kindergarten teachers report a lower frequency of teaching advanced mathematic skills and fewer interactive mathematics activities when there is a larger percentage of children performing below grade level in their class (Gottfried & Kirksey, 2019). Taken together, the characteristics of the peer group with whom individual children share the classroom (i.e., classroom composition) matter for their learning.

Classroom Skill Compositions and its Within- and Cross-domain Effects

One critical component of classroom compositions is classmates' skills. It is clear that there is a great deal of variation in children's preschool-entry skills across domains, due in part to both program enrollment policies and children's varied prior learning experiences (Duncan & Murnane, 2011; McWayne et al., 2004). Put another way, the concentration and distribution of children with certain levels of skills in early academic and executive function domains vary from classroom to classroom. For example, in Head Start classrooms approximately four out of 10 children enter school at risk for being behind in key areas of achievement, with mathematics showing relatively higher developmental risk (Halle et al., 2012). Accordingly, differences in classmates' skills are a key feature of classrooms that may explain individual children's skills gains, and the pattern of these associations may vary by skill domain.

Average Skill of Classmates

Most commonly, researchers examining the role of classmates' skills for children's development in the preschool context do so by estimating the associations between classmates' average level of a skill within a domain at school entry and individual children's gains across the school year in the same domain (i.e., within-domain peer effects; Atkins-Burnett et al., 2017; Foster et al., 2020; Justice et al., 2011; Mashburn et al., 2009; Ribeiro et al., 2017). These studies primarily focus on the role of peers' language skills and suggest that being with classmates who demonstrate higher language skills in the fall results in greater growth in individual preschoolers' expressive and receptive vocabulary (Atkins-Burnett et al., 2017; Foster et al., 2020; Mashburn et al., 2009). One study focusing on the role of preschool classmates' average levels of self-regulation also revealed similar benefits of being with more self-regulated peers (Montroy et al., 2016).

However, not all studies have detected a significant benefit of higher average classmates' skills within preschool classrooms (Justice et al., 2011; Kohl et al., 2021; Ribeiro et al., 2017; Weiland & Yoshikawa, 2014), and effect sizes vary considerably across studies, with estimates ranging from .02 to .36 (Atkins-Burnett et al., 2017; Foster et al., 2020; Henry & Rickman, 2007; Mashburn et al., 2009). These inconsistent patterns may be due to several factors. First, older studies sampled an average of four to seven children per classroom to capture classmates' skills, which represents less than half of each classroom's enrollment (Henry & Rickman, 2007; Justice et al., 2011; Mashburn et al., 2009). The above is of note because the small number of sampled children per classroom may not accurately capture the distribution of classmates' skills. Second, studies also differ in terms of the population their sample represented, ranging from samples drawn from outside the United States (Kohl et al., 2021; Ribeiro et al., 2017) to samples selected from a wider geographic region (Foster et al., 2020; Mashburn et al., 2009) or a single state in the United States (Atkins-Burnett et al., 2017; Henry & Rickman, 2007; Justice et al., 2011). The distinct samples combined with the varied analytic approaches used in previous studies limit our ability to draw comparisons and conclusions across studies. Therefore, continued work is needed to examine whether findings replicate across different samples using parallel analytic methods.

In addition to within-domain peer effects, there is also an emerging body of research that has examined how peers' average levels of a skill in one domain link to individual preschoolers' gains in another domain (i.e., cross-domain peer effects; Montroy et al., 2016; Skibbe et al., 2012). This literature has emerged from evidence that children's own skills are interrelated (Fuhs et al., 2014; Schmitt et al., 2017). For example, children with better-developed executive function were more likely to engage in academic learning activities, which leads to more gains in language, literacy, and mathematics (Bohlmann & Downer, 2016; Nesbitt et al., 2015). Higher levels of language and literacy skills also allow children to make better meaning of mathematics problems (Purpura et al., 2011; Vukovic & Lesaux, 2013) and learn others' expectations and rules for well-regulated behaviors (Vallotton & Ayoub, 2011).

Given the above, classmates' language, literacy, math, and executive function skills may have cross-domain associations with individual preschoolers' early academic skills through either direct or indirect pathways. A low average level of language and literacy skills across the classroom may limit the chance for children to exchange thoughts with peers regarding problem-solving strategies and social rules (Lin et al., 2016), and increase challenges for the teacher to introduce cognitively-stimulating academic content and communicate classroom behavioral expectations (Rjosk et al., 2014; Sawyer et al., 2018). As for executive function skills, a higher concentration of peers with lower levels of behavioral regulation may limit children's opportunities to model peers' self-regulated learning and require more time spent on class management instead of academic instruction, thereby resulting in less optimal growth in children's development (Montroy et al., 2016; Skibbe et al., 2012).

Heterogeneity in Classmates' Skills

In addition to average levels of classmates' skills, another aspect of skill compositions that has garnered much less attention is the heterogeneity of classmates' skills, especially in the preschool context. The heterogeneity of peers' skills is commonly operationalized as the standard deviation of classmates' skills, representing the degree of individual classmates' skills deviating from the classroom mean. There has been increasing interest in whether heterogeneous or homogenous classrooms are more beneficial to children's development in the elementary school context; however, both theoretical and empirical evidence is mixed in this regard (Ben-Ari & Kedem-Friedrich, 2000; Hanushek et al., 2003; Kuzmina & Ivanova, 2018).

From a social cognitive perspective, learning in a heterogeneous classroom may foster children's early academic and executive function skills because children may have a variety of learning opportunities, such as confronting and exchanging different ideas and reinforcing their own knowledge and communicative skills through mutual discussion (Ben-Ari & Kedem-Friedrich, 2000). Despite these theoretical assertions regarding the benefits of heterogeneous classrooms, some scholars argue that it may be easier for a teacher to tailor instruction and learning materials to meet children's needs and interests if their skills are more homogenous (Tomlinson et al., 2003). Supporting these different arguments, studies investigating the consequences of classmates' skill heterogeneity during the elementary school years have also provided inconsistent evidence. Whereas some studies suggest the advantages of having classmates with more heterogeneous skills for children's language and literacy development (Kuzmina & Ivanova, 2018), other studies examining classroom ability groupings indicate that grouping children based on their skills within classrooms or schools benefit elementary students' general learning behaviors and achievement, which supports the merits of homogenous classrooms (Duflo et al., 2011; Hong et al., 2012). There are also studies that document null associations between classmates' skill heterogeneity and elementary students' gains in academic achievement and executive function (Finch et al., 2019; Hanushek et al., 2003).

These incongruent findings require continued investigation in order to inform classroom placement practices and policies. Importantly, previous studies primarily focused on elementary school children (Duflo et al., 2011; Finch et al., 2019; Hong et al., 2012; Hanushek et al., 2003; Kuzmina & Ivanova, 2018), which leaves a knowledge gap in preschool settings. Classmates' skill heterogeneity may operate differently for preschoolers as they may be less likely to engage in constructive peer-peer interactions and knowledge exchanges to leverage the learning opportunities of being in a heterogenous group (Sills et al., 2016). However, the larger skill heterogeneity within classrooms may pose greater challenges, particularly for preschool teachers, because they experience similar demands to adjust instruction in whole group activities (approximately 40% of the day; Pianta et al., 2018) whereas few preschool curricula offer guidelines to help teachers to individualize their instruction (Skibbe et al., 2016). Alternatively, the relatively larger proportion of time spent in small group activities and center play may make it easier for teachers to indivdiualize instruction in the preschool context (Justice et al., 2022).

Peers Matter Differentially for Some Subgroups of Children

Children enter classrooms with a wide range of school-entry skills and sociocultural experiences that can either facilitate or hinder the extent to which they benefit from peer experiences within the classroom. Therefore, previous studies focus on the role of different aspects of children's backgrounds and previous experiences in differentiating peer effects, including children's preschool-entry skills, maternal education, household income, and home language experiences (Atkins-Burnett et al., 2017; Justice et al., 2011; Mashburn et al., 2009; Ribeiro et al., 2017; Schmerse, 2021). However, these studies also reveal different patterns in terms of how peer effects matter for different subgroups of children. One of the common patterns corresponds to a *compensatory effect*, whereby exposure to more skilled peers may compensate for children's disadvantages in school-entry skills and socioeconomic or linguistical background (Atkins-Burnett et al., 2017; Hanushek et al., 2003; Justice et al., 2011, 2014; Ribeiro et al., 2017; Schmerse, 2021). For example, Justice and colleagues (2011) found benefits of being with more skilled peers, but only for the language gains of children who had lower school entry skills. In contrast, fewer studies find support for the *Matthew effect*, whereby children with higher skills (Mashburn et al., 2009) or from more advantaged backgrounds (Luyten et al., 2009) who possess more resources make better use of learning opportunities from being with more skilled peers. Notably, there are also studies that find no support for either pattern in terms of school-entry skills (Kohl et al., 2021; Ribeiro et al., 2017), home language (Kohl et al., 2021), or children's socioeconomic background (Kohl et al., 2021).

With that said, most of the above-reviewed studies with preschoolers focused on whether children's characteristics and family background moderate the effect of classmates' mean language skills (Justice et al., 2011; Mashburn et al., 2009; Ribeiro et al., 2017). As such, we know little about whether these associations emerge for other domains of development and whether the effects of classmates' skill heterogeneity differ for different subgroups of children in the preschool context. Initial evidence on primary school students suggests that the positive association between classmates' skill heterogeneity in language and literacy skills and individual children's learning is stronger for children with relatively low initial skills (Kuzmina & Ivanova, 2018). This suggests that children with relatively low initial skills may benefit more from diversely-skilled peers and potentially more varied instructional content of heterogenous classrooms than their higher-skilled classmates (Kuzmina & Ivanova, 2018). However, studies examining the effect of homogeneous grouping suggest that kindergarteners with low to average abilities showed greater literacy growth when they were frequently placed in small groups with similar abilities and experienced adequate instructional time (Hong et al., 2012).

The Current Study

The extant literature suggests that both the mean and heterogeneity in the skills of classmates may be associated with preschoolers' early academic and executive function outcomes and that these associations may differ for children with different levels of initial skills and sociocultural backgrounds. Even though research to date has been critical in advancing our knowledge regarding peer effects and the role of classroom skill compositions, there are several important ways in which this study builds on prior research including focusing on the role of heterogeneity in peer skills and peer effects among skills beyond just language, cross-domain associations among these skills, and the variation of peer effects for different subgroups of children in the preschool context. Specifically, this study aims to further our understanding by examining: 1) To what extent are classmates' skill mean and heterogeneity associated with preschoolers' skill gains in the areas of language, literacy, mathematics, and executive function? and 2) To what extent do the above associations vary depending on children's preschool-entry skills, maternal education, household income, and home language? Due to the sparse and inconsistent evidence, we address these questions in an exploratory manner.

Given the inconsistent findings and effect sizes on peer effects documented in previous studies, we utilize two large and contemporary datasets to examine whether the associations of interest can be replicated in different populations and reconcile some of the discrepancies reported in previous studies. One dataset draws from families and children in the state of Virginia who are enrolled in a mixed-delivery and publicly funded preschool program targeted at low-income families. These data include a substantial number of immigrants and non-native English speakers who represent the fastest growing segment of the school-age population (Ansari & Crosnoe, 2018). The other dataset, the Head Start Family and Child Experiences Survey 2014 (FACES), is nationally representative of children and families enrolled in the largest federally funded preschool program in the United States, Head Start. Both datasets include a larger number of children per classroom (71% and 50% of each classroom's total enrollment, respectively for the Virginia dataset and FACES; more than Atkins-Burnett et al., 2017; Henry & Rickman, 2007; Justice et al., 2011; Mashburn et al., 2009; Ribeiro et al., 2017), which allow us to capture the variation in classmates' skills within and across classrooms. In addition, each dataset represents different groups of racially/ethnically diverse and low-income populations, which enables a more holistic snapshot of children and their experiences in publicly funded preschool programs and a test for the variation in peer effects for subgroups of children. Moreover, to avoid masking the unique characteristics of each population, we estimate the associations with interest separately with the two datasets instead of combining them. The similarities and/or differences across these datasets in terms of the sampling strategies and the characteristics of sampled children, classrooms, and teachers can provide complementary perspectives regarding the associations of interest using a parallel analytic framework. In doing so, the present study provides a novel opportunity to compare the manifestation of peer effects in different populations and test the generalizability of peer effects in early childhood education systems in the United States.

Method

Data Sources and Procedure

Data used for this study were drawn from two different sources. The first dataset was drawn from a large, culturally, and linguistically diverse county located in Virginia in the 2016 - 2017 school year. This county blends local, state, and federal funding to provide low-cost or free public preschool education to over 2000 eligible low-income families. Notably, 18% of the population in this county were new to the United States, and 53% of children in kindergarten came from families speaking a language other than English as their primary language. The Virginia study recruited preschool teachers from the entire population of preschool classrooms with different combinations of funding sources. All teachers in Head Start and state-funded public school programs were eligible and teachers in community-based private centers were eligible if the center enrolled at least five publicly funded pre-K children. Researchers first sent a flyer with a description of the project to the center director. After center directors indicated an interest in participating, 156 teachers were contacted to describe the project in more detail and to obtain consent forms. Overall, a total of 138 teachers from 83 different schools/centers who met eligibility requirements returned consent forms to the research team and were considered enrolled in the study (88% consent rate; range of classrooms per school/center = 1-9; roughly 1.66 classrooms per school/center). Within the consented classrooms, child participation consent was received from 80% of parents and guardians who had children eligible for participation. Children were eligible to participate if they were enrolled in the pre-K program, turned four years of age by the start of the study, and were not receiving special education services (except for speech). These sampling criteria resulted in a final sample of 1,498 children, for an average of 10.86 (range = 1-18) children from each classroom (i.e., 71% of each classroom's enrollment, range = 6% - 100%).

The second dataset was the FACES 2014, a nationally representative sample of Head Start children and families in 2014 – 2015. FACES followed a multistage sample design to ensure a representative national sample. In the first two stages (Head Start program and centers within programs), FACES sampled with a probability proportional to size design and obtained a total of 60 programs and 118 centers. Then, one or two classrooms within a center and 12 children within a classroom were selected with equal probability. Consent was received from 93% of eligible parents, resulting in a total of 2,462 children across 247 classrooms in the fall of 2014. This sample size decreased to 2,206 children across 245 classrooms in the spring of 2015 due to attrition, which corresponds to 9.00 (range = 1-18) children per classroom (i.e., 50% of classroom enrollment, range = 5% - 100%). Child assessments were conducted in both the fall of 2014 and the spring of 2015. Family background information was collected in the fall of 2014 and teacher surveys were completed in the spring of 2015.

Participants

To maximize the reliability of the measure of classroom skill composition and ensure that each focal child's measure of classmates' skill included information from at least three classmates that was available in each dataset, we excluded classrooms that had less than four children with child assessments (15 classrooms in the Virginia dataset and 10 classrooms in the FACES). This criteria resulted in a subsample of 1,467 children (M_{age} = 4.45 years, SD_{age} = 0.29) across 123 preschool classrooms drawn from the Virginia dataset. For FACES, we also excluded children who had no valid longitudinal weight that is required to draw nationally representative inferences and account for the cross-wave attrition, which resulted in 1,711 out of 2,206 initially sampled children (M_{age} = 4.00 years, SD_{age} = 0.57) across 207 classrooms. Several important similarities and differences exist across the two study samples. First, the classrooms in the Virginia data were diverse in terms of auspice/funding (Head Start, 23%; Private Center, 15%; Public School, 62%), whereas FACES included 60 Head Start programs in the 10 delegated regions across the United States. Second, 79% of children in the Virginia data spoke languages other than English at home whereas in FACES, 76% of children spoke English at home. Third, both datasets were diverse in terms of children's ethnicity/race (Hispanic, 59%/41%; Black, 17%/24%; White, 10%/26%; Asian/other race, 14%/9% for the Virginia data and FACES, respectively); however, the Virginia data included a larger proportion of non-White children. Fourth, children in the Virginia data were 5 months older and were less diverse in age than those sampled in FACES, which included both 3- and 4-year-olds. Fifth, both datasets represent low-income populations (Virginia, M = 0.86, SD = .54; FACES, M = 0.91, SD = .77). Sixth, average years of education among mothers in both datasets was slightly above a high school education; however, parents in the Virginia data were 5 years older than parents in FACES. Seventh, teachers in both datasets were diverse (White, 58%/43%; Black, 21%/28%; Hispanic, 11%/19%; Asian/Other, 10%/11%), but there were more non-White teachers in FACES. Lastly, teachers had over 14 years of teaching experience and over 70% of teachers had obtained a Bachelor's Degree or higher across both samples; however, teachers in Virginia were more educated and experienced. More detailed demographic information can be found in Table 1.

Measures

Children's Early Academic and Executive Functioning Skills

Language Skills. In the Virginia dataset, children's language skills were assessed with the Picture Vocabulary subtest of the Woodcock-Johnson III (α = .81; WJ-III; Woodcock, 2001), which required children to name the depicted objects in a series of pictures. In FACES, children were assessed with the Peabody Picture Vocabulary Test 3rd edition (α = .97; Dunn & Dunn, 1997), which asked children to point to a picture from the four pictures shown on each card. For the purposes of the present study, the W-scores, rather than the raw or standardized scores, were used. The W-score is a special transformation of the Rasch ability scale, which is centered on 500 (approximation of the average score of a 10-year-old child). The typical range of W abilities within a test is about 430 to 550 and varies depending on the trait being measured (McGrew & Woodcock, 1989). The W-score is more ideal as a measure of classroom skill compositions because 1) it is an indicator of children's absolute skills instead of relative skills compared with age norms, and 2) it allows for interpretation of growth in specific skills over time (Jaffe, 2009).

Literacy Skills. Across both datasets, children's literacy skills (W-score) were assessed with the Letter-Word Identification subtest of the WJ-III, which required children to distinguish and pronounce letters and words in a book (α = .94; Woodcock, 2001). Note that the WJ-III norms used in FACES and Virginia datasets were based on Census population projection for 2000 and 2005, respectively, which do not allow for a comparison of children's skill scores across the two datasets.

Table 1
Demographic information for sample in the Virginia data (n= 1,467 children from 123 classrooms) and weighted sample in FACES (n= 1,711 children from 207 classrooms)

	Virginia data *M (SD)*	FACES *M (SD)*
Child and family characteristics		
Child gender: Female (%)	740 (49.76)	858 (50.20)
Child race/ethnicity		
White	152 (10.38)	448 (26.23)
Black	249 (17.00)	406 (23.76)
Hispanic	859 (58.63)	704 (41.18)
Asian/other	205 (13.99)	151 (8.84)
Primary language spoken at home: English (%)	309 (20.82)	1,299 (75.94)
Child age (months)	53.40 (3.48)	48.11 (6.79)
Parent age (years)	34.15 (7.16)	29.29 (6.06)
Parent years of education	12.63 (1.80)	12.38 (1.30)
Income-to-needs	0.86 (0.54)	0.91 (0.77)
Number of adults	2.28 (1.06)	1.89 (.82)
Number of children	2.50 (1.25)	1.33 (1.24)
Teacher and classroom characteristics		
Preschool arrangement type (%)		
Head Start	32 (23.19)	207 (100.00)
Private center	21 (15.22)	0 (0.00)
Public school	85 (61.59)	0 (0.00)
Teacher race/ethnicity (%)		
White	65 (58.04)	87 (42.59)
Black	24 (21.43)	56 (27.57)
Hispanic	12 (10.71)	39 (19.01)
Asian and other	11 (9.82)	23 (10.82)
Teacher early childhood major: Yes (%)	44 (39.29)	190 (93.35)
Teacher years of education	16.86 (1.60)	15.77 (1.46)
Teacher age	43.14 (11.59)	44.16 (11.38)
Teacher teaching years	15.68 (9.73)	14.58 (9.32)
Class size	16.86 (1.85)	17.36 (2.11)
Classroom income-to-need	0.88 (0.30)	0.91 (0.37)
Percentage of boys	.51 (.11)	.44 (.16)
Percentage of IEP	.08 (.08)	.06 (.10)
Percentage of non-white	.77(.17)	.68 (.37)
N	1,467	1,711

Note. IEP = children with Individualized Education Plan

Math Skills. To assess children's math skills (W-score), both the Virginia data and FACES used the Applied Problems subtest of the WJ-III (α = .93; Woodcock, 2001). This subset included increasingly difficult math problems and required children to identify the operation and perform basic calculations based on given information in the problem.

Executive Functioning. We focused on one component of executive function, inhibitory control (Blair, 2016). This focus is due to the associations between higher levels of inhibitory control and classroom behaviors, including active participation in learning activities, positive social-learning interactions, and fewer disruptive behaviors (Nelson et al., 2017; Nesbitt et al., 2015). The Pencil Tap Test was used to assess children's inhibitory control in both datasets (Smith-Donald et al., 2007). As part of the assessment, children were instructed to inhibit their dominant response and display the required response by tapping once when the assessor taps twice and vice visa. To complete this test, children were required to sustain or shift attention to retain the instructed information and handle conflict stimuli and control the irrelevant or incorrect responses. Percent of correct responses out of 16 trials was used as the final score.

Classroom Skill Compositions

Informed by previous studies, we used the leave-out method to calculate classmates' skill means and heterogeneity in the fall so that children's own scores were not included in the calculation of classrooms' skill compositions (Finch et al., 2019; Mashburn et al., 2009). The leave-out means of assessed classmates of a targeted child scores ($N-1$) on the above-listed assessments in the fall were calculated to represent their classmates' average levels of language, literacy, math, and executive function that the targeted child was exposed to. The leave-out standard deviations of assessed classmates of a focal child scores ($N-1$) on those assessments in the fall were calculated for each focal child to represent the degree of heterogeneity in their classmates' skills that the child experienced. The mean, standard deviation, and range of classmates' skill mean and heterogeneity across two datasets are shown in Table 2.

Despite the income eligibility of programs under study, it is important to note that there was a great deal of variability in classroom skill composition. To illustrate the variability between classrooms, we divided the standard deviation of classmates' skill mean by the difference between fall and spring for a given skill. Then, we multiplied the generated estimate by the time between fall and spring assessments, which mapped the variability in skill compositions between classrooms with months of development. Doing so revealed that the differences in classrooms' skill compositions between classrooms in the fall corresponds to 4~6 months (Virginia dataset) and 7~9 months (FACES) in performance. To capture the variability in classroom skill compositions within classrooms, we divided the mean of classmates' skill heterogeneity by the difference between fall and spring for a given skill, which then was multiplied by the average time lapse. In doing so, we found that the variability in classmates' skill within classrooms equals 7~10 months (Virginia dataset) and 9~15 months (FACES). These calculations suggest that 1) there is variability in classroom skill compositions both within and between classrooms; 2) the variability within classrooms is larger than between classrooms; and 3) FACES has larger variability than the Virginia dataset.

Covariates

Previous studies suggest that the estimation of the outcomes of classroom skill compositions are sensitive to endogenous factors, which influence both children's classroom enrollment and experiences along with their skills gains (Fletcher, 2010). To reduce estimation bias due to issues of endogeneity, we included a set of child- and classroom-level covariates that may associate with children's enrollment in different types of classrooms and children's skill gains (Aguiar & Aguiar, 2020; Bassok et al., 2018; Crosnoe et al., 2016; Eslava et al., 2016; Pianta et al., 2005). At the child level, we controlled for *child characteristics* (parent-reported child age, gender, race/ethnicity), *household resources* (home language, parent age, years of education, household income-to-needs), and *household structure* (the number of adults and children in the home). At the classroom level, we controlled for *teacher characteristics* (teacher's age, race, years of education, teaching years, whether the teacher has an early childhood major or not), and *other classroom composition indicators* (class size, average classroom income-to-needs, the percentage of boys, the percentage of non-white children, and the percentage of children with special needs). Given the different types of classrooms sampled in the Virginia data (i.e., Head Start, private center, public school), classroom type was included as an additional control in analyses of Virginia (but not FACES) data. In addition, we adjusted for the time difference between pre-and post-test. The correlations between classroom-level covariates and classroom skill compositions are presented in Supplemental Tables 1a and 1b.

Analytic Approach

To begin, we determined the missingness for both datasets. In the Virginia data, the Pencil Tap Test assessing children's executive function in fall had the highest rate of missingness (25%), followed by family income-to-need ratio (22%) and parents' education level (20%). In FACES, the Pencil Tap Test in fall had the highest rate of missingness (41%), followed by the Pencial Tap Test in the spring (20%) and literacy assessment in the fall (18%). Little's MCAR test suggested that missing data were not missing completely at random for both datasets (χ^2 (2630) = 4635.09, p < 0.001 for the Virginia data; χ^2 (2884) = 5434.15, p < 0.001 for FACES). Further exploration of the missingness patterns indicated a high possibility of missing at random.

Table 2
Descriptives for dependent variables, independent variables, and moderators for the Virginia data and weighted descriptives for FACES.

	Virginia data		FACES	
Variables	Mean(*SD*)	Range	Mean(*SD*)	Range
Individual child's skills in spring				
Language: WJ14 Picture Vocabulary/PPVT	453.66(16.32)	374-491	113.15(19.27)	43-168
Literacy: WJ1 Letter-Word	341.26(28.18)	264-519	328.58(29.58)	264-519
Mathematics: WJ10 Applied Problem	403.28(23.54)	318-453	391.15 (28.26)	318-458
Executive function: Pencil Tap Test	.71(.31)	0-1	.59(.35)	0-1
Individual child's skills in fall				
Language: WJ14 Picture Vocabulary/PPVT	440.76(24.60)	374-481	105.66(19.82)	43-157
Literacy: WJ1 Letter-Word	317.05(29.95)	264-489	314.71(27.41)	264-486
Mathematics: WJ10 Applied Problem	380.55(31.90)	318-449	377.90(31.92)	318-453
Executive function: Pencil Tap Test	.49(.35)	0-1	.47(.35)	0-1
Classmates' skill means				
Language mean	442.45(11.35)	401.89-466.86	104.84(12.09)	62.50-136.20
Literacy mean	318.59(12.12)	287.00-361.40	313.87(15.07)	270.00-388.67
Mathematics mean	382.01(12.92)	349.00-425.50	377.30(18.48)	318.00-436.00
Executive function mean	.50(.14)	.13-.83	.45(.19)	.00 -1.00
Classmates' skill heterogeneity				
Language *SD*	20.82(8.37)	5.18-45.25	16.20(5.63)	1.41-36.08
Literacy *SD*	27.47(7.85)	5.95-50.31	23.03(8.93)	0.00-70.57
Mathematics *SD*	29.57(7.94)	6.60-45.71	27.43(9.31)	0.00-55.15
Executive function *SD*	.34(.06)	.13-.48	.32(.11)	.00-.62

In the Virginia data, missingness of children's post-test scores were correlated with child age, race/ethnicity, and parent's education level; in FACES, children's post-test scores were shown to be missing as a function of child age, home language, parent's education level, and teacher's race. To minimize the potentially biased estimation and the reduction in statistical power by analyzing complete case data (Enders, 2013), we imputed 50 data sets with chained equations. All the analyzed variables in each dataset were included in the imputation model and both observed and imputed values were used to estimate missing values.

The analyses across two datasets were performed using a two-level multilevel modeling (MLM) framework to account for the nesting of children in classrooms, with slight differences in specifications due to the different sampling strategies. In the Virginia data, we controlled for preschool arrangement type (i.e., Head Start, private center, and public school). In comparison, because FACES adopted a multistage sampling strategy, we incorporated a child-level longitudinal weight (PRA12OCW) and classroom-level weight (T2CLSWT) in our MLM models to adjust for the differential probabilities of selection and sample attrition at both levels and reduce estimation bias. We also incorporated the effective scaling method in MLM models to estimate the child-level variance more accurately (Shen & Konstantopoulos, 2022). In addition, because the selection of classrooms were nested within primary sampling units (PSUs), we also used cluster-robust standard errors at the PSU level. Across both datasets, all continuous variables were standardized to have a mean of 0 and a standard deviation of 1; thus, the estimates for both datasets can be interpreted as standardized coefficients.

To establish the associations between classroom skill compositions (i.e., classmates' skill mean and skill heterogeneity) and children's skill gains, we began by estimating a series of models that included children's skills in the spring regressed on classroom skill compositions, separately for each outcome domain. Classroom skill compositions were estimated at level 1 controlling for children's pretest score and other child covariates at level 1 and classroom covariates at level 2. Given the high correlations between classmates' skill mean and skill heterogeneity for some of the assessed domains (see Table 3a and 3b for each dataset), we included each classmates' skill mean and each skill heterogeneity variable in separate models. In each dataset, a total of 32 models were estimated to test the main effects of the four classmates' skill mean variables and the four heterogeneity variables on the four skill outcomes. An example of the equation for models that tested the main effects of classmates' language mean is presented below:

Level 1: $Y_{ij} = \beta_{0j} + \beta_{1j}(classmates'\ language\ mean)_{ij} + \beta_{2j}(pretest\ score)_{ij} + \beta_{3j}(Child\ covariates)_{ij} + e_{ij}$

Level 2: $\beta_{0j} = \gamma_{00} + \gamma_{01}(classroom\ covariates)_j + \mu_{0j}$

To explore the moderating role of children's initial skills, household income, maternal education, and home language, we estimated another series of models in which the level 1 equation was changed by adding an interaction term of these moderators and skill compositions in separate models and the level 2 equation remained the same. In each dataset, a total of 128 models were estimated to test if the above 32 main associations vary by the four moderators. All models adjusted for children's initial skill in the fall corresponding to the domain of outcome skill as well as the child- and classroom-level covariates mentioned in the measures section (see Table 1). For example, the level 1 equation of a model testing the moderating effect of children's home language in the associations between classmates' language mean and outcomes is:

Level 1: $Y_{ij} = \beta_{0j} + \beta_{1j}(classmates'language\ mean)_{ij} + \beta_{2j}(pretest\ score)_{ij} + \beta_{3j}(Child\ covariates)_{ij} + \beta_{4j}(classmates'\ language\ mean*children's\ home\ language)_{ij} + e_{ij}$

As a sensitivity check, we reestimated all models using the whole sample, which included classrooms with less than four children with direct assessments. And in instances in which only the study child was sampled, we imputed classmates' skill data. Despite slight changes in the effect sizes, the results were substantively the same (see Supplemental Tables 2 and 3).We conducted exploratory analyses with both the mean and heterogeneity of classmates' skills in each domain in the same model. The coefficients of classmates' skill mean and heterogeneity remained largely the same (see Supplemental Table 4). Lastly, to ensure our results were robust to the entering of child- and classroom-level covariates, we reran all the models with child-level covariates and classroom-level covariates entered sequentially. The results were not sensitive to the sequencing of covariates (see Supplemental Tables 5 and 6).

Results

Classroom Skill Compositions and Children's Outcomes

We began by testing the within- and across-domain associations of classmates' language, literacy, math, and executive function mean skills with individual children's spring scores in these domains. In the Virginia dataset, none of the within-domain associations were significant after

Table 3a
Correlations between skill means and skill heterogeneity for the Virginia data

	Mean				Heterogeneity			
	1. Language	2. Literacy	3. Math	4. EF	5. Language	6. Literacy	7. Math	8. EF
1.	-							
2.	.69***	-						
3.	.76***	.69***	-					
4.	.35***	.43***	.51***	-				
5.	-.71***	-.45***	-.59***	-.18***	-			
6.	.04	.27***	.11***	.29***	.17***	-		
7.	-.41***	-.37***	-.60***	-.07**	.53***	.21***	-	
8.	.11**	.06	.09**	.09**	.01	.29***	.11**	-

*p<0.05, **p<0.01, ***p<0.001.

Table 3b
Weighted correlations between skill means and skill heterogeneity for the weighted FACES

	Mean				Heterogeneity			
	1. Language	2. Literacy	3. Math	4. EF	5. Language	6. Literacy	7. Math	8. EF
1.	-							
2.	.45***	-						
3.	.67***	.62***	-					
4.	.53***	.26***	.50***	-				
5.	-.04	-.02	.12***	.04	-			
6.	.12***	.24***	.10***	.08**	.30***	-		
7.	.01	-.10***	-.23***	.01	.32***	.26***	-	
8.	.22***	.11***	.18***	.39***	.14***	.17***	.18***	-

*p<0.05, **p<0.01, ***p<0.001.

controlling for the fall scores in the corresponding domain and all other covariates (β = |.00 - .04|; see Table 4). Further examination of cross-domain associations also revealed largely null patterns (β = |.00 – .05|; see Table 4). In FACES, the four within-domain effect sizes of classmates' skill mean on the corresponding skills are somewhat larger as compared with the Virginia dataset as shown in Table 4 (β = .03 - .09), including three significant within-domain associations between classmates' mean literacy skills and children's spring literacy skills, between classmates' mean math skills and children's spring math skills, and between classmates' mean EF and children's spring EF (β = .07, SE = .03, p < .05; β = .06, SE = .02, p < .05; β = .09, SE = .04, p < .05). The cross-domain associations between skill means were also somewhat larger in FACES (β = |.01 - .15|; see Table 4), but only two significant associations were documented, including classmates' mean language and children's spring executive function (β = .15, SE = .04, p < .01) and between classmates' mean literacy and children's spring math (β = .07, SE = .02, p < .05). Taken together, approximately one third of the models revealed significant effects of classmates' skill mean in the FACES data; however, only one out of sixteen were significant in the Virginia data.

Next, we tested the within- and across-domain associations between classmates' skill heterogeneity across language, literacy, math, and executive function and individual children's spring scores in these domains in separate models. Results of both datasets revealed largely null patterns for the classmates' skill heterogeneity within-domain associations (β = |.01 - .02| for the Virginia data; |.01 - .02| for FACES see Table 5). Additionally, none of the cross-domain associations were significant in the Virginia data (β = |.00 - .04|; see Table 5). Cross-domain associations were also largely nonsignificant in FACES (β = |.00 - .07|), except for the one significant association between classmates' EF heterogeneity and children's spring literacy skills (β =.07, SE = .03, p < .05; see Table 5).

Child-Level Variability in the Outcomes of Classroom Skill Compositions

Having established the associations between classroom skill compositions and children's early learning, we next tested whether children's initial skills, household income, maternal education, and home language moderated the above associations in separate models. For classmates' mean skills, only five out of sixty-four interactions tested in the Virginia data were statistically significant and four out of the sixty-four interactions in FACES were statistically significant (see Table 4). In terms of variability in associations between classmates' skill heterogeneity and children's early learning, only two significant interactions were documented in the Virginia data and only one significant interaction was documented in FACES (see Table 5). Taken together, there were *no* consistent patterns of moderating effects within and across the two datasets, providing little evidence that the associations between classroom skill compositions and children's spring skills vary as a function of the examined moderators.

Discussion

Understanding how preschool experiences shape children's early academic development and executive functioning is critical, as these skills shape children's long-term school success (Duncan et al., 2007) and have lifelong consequences for health and economic well-being (Reynolds et al., 2011). Given the substantial time children spend with classmates, there has been ongoing interest in understanding the role of classmate skills in preschoolers' learning (e.g., Justice et al., 2011; Kohl et al., 2021; Mashburn et al., 2009). Using two large samples with a considerable representation of children in the classroom, our work adds to previous studies by focusing on two aspects of classroom skill compositions, namely skill mean and skill heterogeneity, and exploring how they are associated with children's development of language, literacy, mathematics, and executive function skills. As part of this effort, we also sought to identify whether the links between classroom skill compositions and children's early learning vary by child and family characteristics. We found that across the two datasets, there was limited evidence for: (a) both within-domain and cross-domain effects of classmates' skill means and skill heterogeneity except for small within-domain associations in FACES; and (b) differential associations for different subgroups of children. We discuss the themes of our work in more detail below.

Table 4
Associations between classmates' skill means, interaction between skill means and child and family characteristics, and children's early academic and executive function outcomes across two datasets using a subsample of classrooms with at least four children

	Virginia data (n = 1,467 children across 123 classrooms)				FACES (n = 1,711 children across 207 classrooms)			
	Language	Literacy	Math	EF	Language	Literacy	Math	EF
Classmates' language mean	-.035(.022)	-.012(.029)	-.024(.024)	-.028(.030)[a]	.030(.025)	.047(.038)	.050(.030)	.147(.042)**
Classmates' literacy mean	-.003(.021)	.038(.028)	-.028(.024)	-.011(.029)[a]	.018(.020)	.072(.032)*	.067(.022)**	.072(.037)
Classmates' math mean	-.020(.021)	.006(.027)	.003(.023)	.012(.027)	-.023(.023)	.020(.032)	.055(.024)*	.072(.039)[a]
Classmates' EF mean	-.001(.020)	-.054(.027)*	-.037(.023)	.002(.029)	.021(.021)	.004(.031)	-.001(.025)[a]	.087(.036)*
Language mean * Initial skill	.009(.017)[a]	.014(.020)	.020(.020)	.030(.025)	.001(.014)	.052(.028)	.008(.023)[b]	-.067(.027)*
Literacy mean * Initial skill	.020(.017)	.034(.020)	.028(.020)	.009(.025)	.006(.016)	.004(.022)	.017(.024)[a]	-.039(.026)
Math mean * Initial skill	-.004(.018)	.033(.021)	-.001(.020)	.015(.024)	-.000(.015)	.012(.023)	.015(.023)	-.086(.031)**
EF mean * Initial skill	-.035(.017)*	.056(.019)**	-.020(.020)	.017(.024)[a]	.005(.018)	-.021(.026)	-.009(.022)[a]	-.045(.038)
Language mean * Household income	-.034(.016)*	.011(.019)	-.018(.019)	-.001(.024)	-.015(.014)	.018(.029)	-.002(.021)[a]	-.013(.021)
Literacy mean * Household income	-.020(.016)	.003(.018)	-.027(.018)	-.002(.023)[a]	.006(.016)	.013(.033)	.021(.020)[a]	.004(.020)
Math mean * Household income	-.034(.015)*	.002(.017)	-.021(.018)	.007(.022)	-.013(.015)	.032(.035)	.000(.019)	.009(.034)
EF mean * Household income	-.041(.016)*	-.005(.018)	-.014(.020)	.008(.024)	-.002(.014)	-.007(.024)	-.035(.034)[b]	-.035(.031)
Language mean * Maternal education	-.028(.020)[b]	.004(.021)	-.029(.023)	.011(.027)[a]	.013(.018)	.020(.020)	.038(.022)[a]	-.012(.027)
Literacy mean * Maternal education	-.015(.018)	.009(.019)	-.023(.021)	.011(.025)	.017(.018)	.000(.020)	.028(.022)[b]	-.015(.031)
Math mean * Maternal education	-.024(.018)	.012(.020)	-.036(.021)	.016(.025)[a]	.020(.016)	.013(.019)	.036(.025)	-.055(.029)
EF mean * Maternal education	-.019(.018)	.019(.020)	-.030(.020)	-.024(.025)[c]	.029(.016)	.001(.021)	.046(.024)	.000(.031)
Language mean * Home language	-.048(.048)	.049(.055)	-.017(.056)	.046(.065)	.020(.039)	.037(.071)	.070(.057)[a]	-.114(.052)*
Literacy mean * Home language	-.025(.039)	.038(.045)	-.002(.047)	.012(.056)[a]	.027(.037)	.037(.048)	.090(.052)	-.028(.062)
Math mean * Home language	-.043(.041)	-.033(.047)	-.052(.048)	.018(.056)	.027(.032)	.094(.056)	.076(.049)[b]	-.050(.065)
EF mean * Home language	-.023(.043)	.017(.048)	-.069(.049)	-.015(.060)	.014(.039)	.011(.064)	.003(.055)	-.137(.063)*

Note. All estimates correspond to standardized betas and those in parentheses correspond to standard errors. Models controlled for children's initial skill corresponding to the domain of outcome skill, child- (Child age, gender, race/ethnicity, primary home language, parent age, maternal years of education, income to needs, number of adults, number of child in household, time difference between pretest and posttest), and classroom-level covariates (Teacher age, teacher race/ethnicity, years of education, teaching years, whether having a ECE major, classroom type, class size, average classroom income-to-needs, percentage of boys, percentage of non-white children, percentage of children with IEP/IFSP in the classroom). Child race/ethnicity and teacher race/ethnicity were coded as dummy variables (0/1) and the White was the reference category. In the Virginia data, classroom type was also coded as dummy variable and Public School was the reference category.
Each classmates' skill mean and each interaction term was entered in separate models. The coefficients of classmates' skill mean were drawn from models that did not include the interaction terms.
The estimates reported herein correspond to standardized coefficient, because all continuous variables were standardized to have a mean of 0 and standard deviation of 1.
Results combined across 50 imputed datasets except for the ones failing to converge.
[a] 49 imputations.
[b] 48 imputations.
[c] 47 imputations.* $p<0.05$, ** $p<0.01$, *** $p<0.001$.

Limited Evidence for Within- and Cross-domain Effects of Classmates' Skill Mean

We did not find consistent associations between classmates' average skill levels and children's development of language, literacy, math, and executive function skills. These largely nonsignificant patterns emerged for both the within- and cross-domain models and across two datasets. When interpreting these findings, it is important to acknowledge that previous studies have provided contradictory evidence. Several studies have revealed a significant within-domain effect of classmates' average language skills, with effect sizes ranging from .02 - .36 (Atkins-Burnett et al., 2017; Foster et al., 2020; Mashburn et al., 2009) and a cross-domain effect of classmates' executive function on literacy skills (.09 - .12; Montroy et al., 2016). However, other studies have failed to detect significant associations between classmates' average language skills and individual children's language development (Justice et al., 2011; Kohl et al., 2021; Ribeiro et al., 2017; Weiland & Yoshikawa, 2014). The current study served as a step to resolve the inconsistencies in the literature by: (a) extending the focus from classmates' average language skills to include children's literacy, mathematics, and executive function skills and (b) examining both within- and cross-domain peer effects with two large and contemporary datasets. The limited significant findings reported in the current study highlight some important issues for further consideration.

First, previous studies suggest that lower family income is associated with children's overall levels of skills (Magnuson & Waldfogel, 2016), and children enrolled in publicly funded programs targeting low-income families such as Head Start enter with relatively lower academic and executive function skills than national norms (Hulsey et al., 2011). Although our descriptives showed that there was considerable variation in classmates' skills both within and between classrooms, especially in FACES, whether this level of variation is sufficient to promote the types of scaffolding required for more optimal levels of peer-to-peer learning remains unclear. Descriptives of the standardized scores for language, literacy, and mathematics in the fall and spring suggest that children in both datasets scored, on average, lower than the national norms, i.e., less than 100 (86.84 – 97.48 in Virginia dataset; 89.40 – 96.00 in FACES). Given the limited evidence reported herein, continued investigation is needed to examine whether peer effects vary by program enrollment policies. But given the widespread use of publicly funded preschool programs to support children from low-income families, our findings suggest a next step for researchers and practitioners is to identify the conditions that enable peer-to-peer learning in such programs. Another potential explanation for the largely null linear associations is the existence of non-linear effects or thresholds of classmates' skills. That is, the effects of classmates' mean skills may be conditional on the presence of a certain number of higher skilled peers in the classroom.

Second, the findings reported herein call for more nuanced measures to capture classroom composition. The measures of classmates' average skills capture the available skills of classmates; however, they may not reflect the actual skills of peers whom children interact with on a regular basis (Chen et al., 2020). Not all classmates are children's playmates, and children do not interact with and learn from classmates in an equally effective manner. On the one hand, children are inclined to

Table 5
Associations between classmates' skill heterogeneity, interaction between skill heterogeneity and child and family characteristics, and children's early academic and executive function outcomes across two datasets with classrooms using a subsample of classrooms with at least four children

	Virginia data (n = 1,467 children across 123 classrooms)				FACES (n = 1,711 children across 207 classrooms)			
	Language	Literacy	Math	EF	Language	Literacy	Math	EF
Classmates' language heterogeneity	-.012(.021)	-.009(.027)	-.002(.022)	-.010(.028)	-.024(.017)	-.002(.028)	-.027(.022)[c]	.002(.033)
Classmates' literacy heterogeneity	-.001(.019)	.008(.024)	.003(.020)	-.044(.025)	.021(.014)	.014(.033)	.010(.020)	.091(.049)
Classmates' math heterogeneity	-.003(.019)	-.032(.025)	.024(.021)	-.036(.026)[a]	.021(.017)	.039(.033)	-.011(.023)[b]	.023(.038)
Classmates' EF heterogeneity	.009(.018)[a]	.027(.023)	-.007(.020)	-.015(.024)	-.004(.019)	.069(.034)*	.011(.030)[a]	-.005(.032)
Language heterogeneity * Initial skill	-.022(.019)	.003(.018)	-.017(.020)	-.016(.023)	.015(.019)	.018(.030)	.033(.025)	.008(.024)
Literacy heterogeneity * Initial skill	-.036(.029)	.005(.027)	.009(.022)	.013(.023)	-.005(.030)	.024(.022)	-.005(.027)[a]	-.002(.016)
Math heterogeneity * Initial skill	-.050(.019)	-.045(.019)	-.051(.024)*	-.004(.024)	.015(.027)	.029(.024)	.016(.025)[b]	.027(.027)
EF heterogeneity * Initial skill	-.041(.025)	-.026(.023)	.012(.025)	.021(.025)	.012(.027)	.008(.034)	.001(.024)	-.046(.028)
Language heterogeneity * Household income	.032(.017)*	-.020(.019)	.004(.019)	.002(.022)	-.002(.016)	.035(.030)	.019(.022)[b]	.044(.036)
Literacy heterogeneity * Household income	.009(.017)	-.010(.019)	-.006(.020)	.003(.025)[b]	.009(.016)	.012(.028)	.024(.021)[a]	.006(.029)
Math heterogeneity * Household income	.014(.017)	-.022(.019)	-.001(.020)	-.002(.023)[a]	.022(.023)	.015(.030)	.021(.030)[a]	.006(.041)
EF heterogeneity * Household income	.006(.017)[a]	.001(.019)	-.012(.020)	-.027(.023)[b]	.017(.015)	.040(.023)	.065(.034)	.050(.041)
Language heterogeneity * Maternal education	.025(.018)	-.009(.020)	.012(.021)	-.010(.026)[b]	.014(.019)	.041(.030)	.033(.023)	.062(.025)*
Literacy heterogeneity * Maternal education	-.008(.017)	.003(.020)	-.001(.020)	-.003(.024)[a]	.000(.021)	.011(.028)	.012(.024)[a]	.022(.030)
Math heterogeneity * Maternal education	.007(.018)	-.019(.020)	.009(.021)	-.022(.024)	.009(.021)	.036(.029)	.000(.022)[a]	.041(.032)
EF heterogeneity * Maternal education	-.025(.019)[a]	.001(.021)	-.013(.022)	-.046(.025)	.019(.018)	.026(.033)	.013(.027)	-.005(.037)
Language heterogeneity * Home language	-.014(.040)	-.048(.045)	-.053(.046)	-.094(.053)[a]	.046(.040)	.089(.051)	.028(.062)	.096(.063)
Literacy heterogeneity * Home language	-.008(.039)	-.043(.045)	-.027(.045)	-.077(.056)[a]	.004(.059)	-.007(.062)	.033(.073)[a]	-.003(.061)
Math heterogeneity * Home language	-.010(.037)	-.032(.042)	-.045(.044)	-.068(.053)	.018(.048)	.038(.061)	-.004 (.058)	.011(.060)
EF heterogeneity * Home language	.035(.039)	-.032(.044)	.021(.046)	-.099(.055)	-.016(.035)	.009(.079)	-.002(.061)[b]	.031(.067)

Note. All estimates correspond to standardized betas and those in parentheses correspond to standard errors. Models controlled for children's initial skill corresponding to the domain of outcome skill, child-(Child age, gender, race/ethnicity, primary home language, parent age, maternal years of education, income to needs, number of adults, number of child in household, time difference between pretest and posttest), and classroom-level covariates (Teacher age, teacher race/ethnicity, years of education, teaching years, whether having a ECE major, classroom type, class size, average classroom income-to-needs, percentage of boys, percentage of non-white children, percentage of children with IEP/IFSP in the classroom). Child race/ethnicity and teacher race/ethnicity were coded as dummy variables (0/1) and the White was the reference category. In the Virginia data, classroom type was also coded as dummy variable and Public School was the reference category.
Each classmates' skill mean and each interaction term was entered in separate models. The coefficients of classmates' skill heterogeneity were drawn from models that did not include the interaction terms.
The estimates reported herein correspond to standardized coefficient, because all continuous variables were standardized to have a mean of 0 and standard deviation of 1.
Results combined across 50 imputed datasets except for the ones failing to converge.
[a] 49 imputations.
[b] 48 imputations.
[c] 47 imputations.*$p<0.05$, **$p<0.01$, ***$p<0.001$.

choose classmates who are similar to them (e.g., skills, gender, family income) as their playmates, and their skills increase or decrease towards the levels of their playmates (DeLay et al., 2018). On the other hand, whether children can benefit from their peers during peer interactions depends on the extent to which children can consciously observe and model peers' behaviors or cooperatively exchange and discuss conflicting perspectives (Rubin et al., 2006). However, some preschoolers may not yet be able to purposefully and positively interact and learn from each other without adult support (Sills et al., 2016). As such, future studies need to use more nuanced measures capturing peer selection and peer interaction.

There was some evidence suggesting the effects of classmates' skill mean in the FACES data especially for the within-domain associations, with a slightly larger average effect size of around 0.06 compared with the Virginia dataset. The inconsistent findings across the two datasets are somewhat similar to the mixed findings seen in previous studies, which calls for further examination of the variation in peer effects within/between different populations and across different contexts. There are several plausible explanations for the inconsistent findings. First, the Virginia dataset sampled a much more linguistically diverse population with 57% of children speaking Spanish and 23% speaking another non-English language at home, compared both with FACES (21%, Spanish; 2%, other non-English languages) and samples in previous studies (Foster et al., 2021; Justice et al., 2011; Mashburn et al., 2009; Montroy et al., 2016). Because part of the transmission of early academic and executive function skills among children relies on verbal communication in a common language (Kuhn et al., 2014; Ramsook et al., 2020), the diversity in children's home-spoken language in the Virginia data may have constrained the exchange of knowledge and skills among classmates. In addition, our skill composition variables were created from classmates' scores on English assessments, which may fail to capture children's exchange of skills using Spanish and other common languages in the Virginia data. Second, the sample of relatively younger and more age diverse children in FACES may enable children to benefit more from interacting with peers in older age groups (Foster et al., 2021). Lastly, the classmates' skill mean in FACES has larger between-classroom variability than the Virginia dataset according to our calculation (7~9 months vs. 4~6 months), which may be more likely to account for the variation in children's skill gains.

Little Evidence for Within- and Cross-domain Effects of Classmates' Skill Heterogeneity

We also found nonsignificant associations between heterogeneity in classmates' early academic and executive function skills and individual children's skills gains. This nonsignificant pattern emerged across most within- and cross-domain models for both datasets, which aligns with previous studies using primary school samples (Hanushek et al., 2003; Finch et al., 2019). To date, the outcomes of classmates' skill heterogeneity have not been adequately examined in the preschool context and studies in elementary school context have generated mixed findings (Duflo et al., 2011; Finch et al., 2019; Hong et al., 2012; Hanushek et al., 2003; Kuzmina & Ivanova, 2018). Thus, our findings contribute to the literature by demonstrating the extent to which preschoolers benefit from heterogeneity in classmates' early academic and executive function skills.

The overall nonsignificant effects of classmates' skill heterogeneity in the current study are inconsistent with some previous studies in elementary school contexts suggesting that the heterogeneity in peers' skills within classrooms/groups may either facilitate or interfere with children's learning (Kuzmina & Ivanova, 2018; Duflo et al., 2011; Hong et al., 2012). The inconsistent findings across our study and previous studies may be explained by the differences in peer interactions and teachers' instruction across contexts. As noted above, the extent to which children are able to leverage the heterogeneity in peers' skills (e..g., learning from more knowledgeable others, reflecting on their thoughts when encountering different perspectives) also depends on with whom they interact and whether they interact in a frequent and effective manner (DeLay et al., 2018). The nonsignificant findings on classmates' skill heterogeneity occur across two samples of preschoolers with slight age differences, which may suggest that peer interactions during the preschool period may not be efficient enough to enable the manifestations of classmates' skill heterogeneity influence (Sills et al., 2016). Moreover, the effects of classmates' skill heterogeneity may be also contingent on teachers' abilities to deal with the heterogeneity and provide individual children with developmentally appropriate instruction (Ansari, 2017; Tomlinson et al., 2003). To better illustrate the influences of classmates' skill heterogeneity in different contexts, future research should use more nuanced measures of peer interaction and teachers' instructional practices.

Little Variability in Peer Effects by Children's Characteristics and Family Background

Lastly, our findings revealed little evidence that classmates' skill mean and heterogeneity have differential associations with children's skills gains depending on children's initial skills, maternal education, household income, and home language across the two datasets. Combined with the limited evidence of main effects, our findings suggest that neither the average level of classmates' skills nor the heterogeneity in classmates' skills has major consequences for preschoolers' of different academic, socioeconomic, or linguistic backgrounds. These largely null findings are surprising given they provide no support for either the *compensatory effect* or the *Matthew effect* that have been highlighted in previous studies (e.g., Justice et al., 201; Mashburn et al., 2009). However, similar null patterns were also found when previous studies examined the role of classmates' mean language skills with a German sample (Kohl et al., 2021).

Even though the largely null findings are somewhat surprising given the theoretically informed variability in peer effects, the lack of differential associations by children's characteristics warrants more replication work and highlights the necessity to continue testing other potential moderators. The nonsignificant moderators in the current study may be due to the fact that teachers are not focused on leveraging the academic strengths of children's classmates. For example, a survey among early elementary teachers suggested that teachers rated separating behavioral problems instead of heterogeneous ability grouping as most important (Kim et al., 2020). Combined with children's tendency to interact with peers like them in terms of demographic characteristics and skills (DeLay et al., 2018), the skills exchange between higher-and lower-skilled children may be limited.

The extent to which peer effects vary by children's characteristics and family backgrounds may depend on whether and how teachers adjust their instructional content and strategies based on the needs of children. For example, children who entered kindergarten with lower skills were found to benefit more from homogenous ability grouping only when their group received ample instruction from teachers (Hong et al., 2012). As another example, Foster and colleagues (2021) documented a significant main effect of classmates' language skills that was stronger for English speakers in a sample of preschoolers from largely low-income families. It is important to note that their sample came from programs that focused on improving teachers' teaching practices and professional development. As such, future studies need to investigate classroom processes (e.g., teachers' instructional practices and grouping strategies), teacher attitudes and beliefs towards heterogeneity as well as teacher and program focus (e.g., focusing on low or high achievers) as potential moderators in distinguishing the consequence of classroom skill compositions.

Limitations and Future Directions

Overall, our findings provide little evidence that, above and beyond children's own preschool-entry skills, their classmates' skills associate with their development of academic and executive function skills. With that said, there are several limitations and future directions that should be acknowledged. First, we do not know whether our findings generalize to universal preschool programs without income eligibility for enrollment. Previous studies sampling children from universal preschool programs reveal significant effects of classmates' skill mean (Atkins-Burnett et al., 2017; Mashburn et al., 2009). Thus, the associations between classroom skill compositions and preschoolers' academic and executive function outcomes and the variation in peer effects among subgroups of children may be more distinct in universal preschool programs due to potentially larger variability in classmates' skills within and across classrooms (Barnett, 2010). With that said, the two datasets each represent a localized public preschool program and a national early education system that serve children from low-income households in the United States, and thus, the present study provides a comprehensive snapshot of how peer effects work in these systems.

Furthermore, linear associations of classmates' skill means and heterogeneity in skills may not tell the whole story about the relations under investigation. Future studies need to consider other ways to represent classroom skill compositions, such as the percentage of children above or below proficiency levels. Threshold effects should also be tested with a more diverse distribution of classmates' skills. Additionally, this study did not disentangle peer effect from peer selection and did not measure peer interaction within classrooms. With only a record of classmates' skills level, this study, like other studies of peer effects (e.g., Henry & Rickman, 2007; Justice et al., 2011; Mashburn et al., 2009), was unable to identify with whom the study children interacted with and how they interacted with each other. Accordingly, future studies should include more nuanced observations of peer interactions to better understand the mechanisms underlying peer effects.

Next, the Pencil-Tap test used in both datasets to measure children's executive function contained a high rate of missingness in the fall. The examination of missingness patterns suggested that children whose test scores were missing were most likely to be those who were younger at preschool entry. Whether this test is suitable to assess preschoolers' executive function at the school entry remains in question. However, this test has been widely used to assess children's inhibitory control and validated in samples of children ranging from 27 to 73 months and from varying socioeconomic and sociocultural contexts (Raver et al., 2012; Smith-Donald et al., 2007). Instead of only recording children's response accuracy in this test, future studies may also consider recording their response speed to better measure executive function and related peer effects. For example, Finch and colleagues (2019) measured both response accuracy and response speed on EF tasks and found the effect of classmates' EF skills was driven by EF-related speed on the tasks instead of accuracy.

Our test of the moderating role of child and family characteristics failed to reveal variability in the focal associations of interest. However, it is plausible that peer effects emerge for other populations or under different classroom circumstances; as such, future studies should consider other potential moderators, such as teachers' expectations and beliefs, instructional content, and classroom grouping strategies. Additionally, we did not account for the multilevel structure of the data in the imputation models, which better matches the multilevel models we estimated. Lastly, the short time frame between pretest and posttest in both datasets may limit our ability to test relatively long-term effects of classmates' skills. The high associations between children's initial skills in the fall and outcome skills in the spring across both datasets (B = .48 -.77 in the Virginia data and .54-.71 in FACES; see Supplemental Table 7) may also contribute to our null findings. That said, there was still a considerable amount of residual variance in outcome skills in the spring that are unexplained by the initial skills in the fall (range = 41% - 77% in the Virginia data and 49% -71% in FACES).

Conclusions

With the above limitations and future directions in mind, this study added to the existing knowledge about peer effects in preschool settings by examining the role of classmates' skill mean and heterogeneity in two large and contemporary datasets in the United States. As both data sampled a large number of children per classroom and each represented different groups of ethnically/racially diverse and low-income populations, our findings generalize to many current publicly funded preschool systems that target children from low-income families. When taken together, results from the current study revealed limited and small significant within- and cross-domain associations between classmates' skill mean and heterogeneity with preschoolers' early academic and executive function outcomes across two datasets. We also found little evidence for variability in peer effects as a function of preschool-entry skills, maternal education, household income, and home language, suggesting *few* differential associations for preschoolers of different academic, socioeconomic, or linguistic backgrounds. However, there was some indication that classmates' skill means were associated with children's skill gains in the FACES data, especially when considering within-domain associations. Given the inconsistent findings in the current study and mixed findings in previous studies, continued attempts at replication within different populations and across different contexts are necessary. Overall, our results suggest that the average and heterogeneity in classmates' skills do not play a major role in preschoolers' learning, and future research should focus on more nuanced interactions within classrooms to understand the ways in which peer interactions shape development.

CRediT authorship contribution statement

Qingqing Yang: Conceptualization, Methodology, Software, Writing – original draft, Writing – review & editing. **Arya Ansari:** Conceptualization, Supervision, Investigation, Resources, Writing – review & editing. **Kelly M. Purtell:** Conceptualization, Supervision, Writing – review & editing. **Robert C. Pianta:** Conceptualization, Investigation, Resources, Writing – review & editing. **Jessica V. Whittaker:** Conceptualization, Investigation, Resources, Writing – review & editing. **Virginia E. Vitiello:** Conceptualization, Investigation, Resources, Writing – review & editing.

Data availability

The authors do not have permission to share data.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ecresq.2023.04.003.

References

Aguiar, A. L., & Aguiar, C. (2020). Classroom composition and quality in early childhood education: A systematic review. *Children and Youth Services Review, 115*, Article 105086. 10.1016/j.childyouth.2020.105086.

Atkins-Burnett, S., Xue, Y., & Aikens, N. (2017). Peer effects on children's expressive vocabulary development using conceptual scoring in linguistically diverse preschools. *Early Education and Development, 28*(7), 901–920. 10.1080/10409289.2017.1295585.

Bandura, A. (1986). Social foundations of thought and action: A social cognitive theory *(pp. xiii, 617)*. Prentice-Hall, Inc.

Barnett, W. S. (2010). Universal and Targeted Approaches to Preschool Education in the United States. *International Journal of Child Care and Education Policy, 4*(1), 1–12. 10.1007/2288-6729-4-1-1.

Bassok, D., Magouirk, P., Markowitz, A. J., & Player, D. (2018). Are there differences in parents' preferences and search processes across preschool types? Evidence from Louisiana. *Early Childhood Research Quarterly, 44*, 43–54. 10.1016/j.ecresq.2018.01.006.

Ben-Ari, R., & Kedem-Friedrich, P. (2000). Restructuring heterogeneous classes for cognitive development: Social interactive perspective. *Instructional Science, 28*(2), 153–167. 10.1023/A:1003806300757.

Blair, C. (2016). Executive function and early childhood education. *Current Opinion in Behavioral Sciences, 10*, 102–107. 10.1016/j.cobeha.2016.05.009.

Bohlmann, N. L., & Downer, J. T. (2016). Self-regulation and task engagement as predictors of emergent language and literacy skills. *Early Education and Development, 27*(1), 18–37. 10.1080/10409289.2015.1046784.

Booren, L. M., Downer, J. T., & Vitiello, V. E. (2012). Observations of children's interactions with teachers, peers, and tasks across preschool classroom activity settings. *Early Education and Development, 23*(4), 517–538. 10.1080/10409289.2010.548767.

Chen, J., Justice, L. M., Tambyraja, S. R., & Sawyer, B. (2020). Exploring the mechanism through which peer effects operate in preschool classrooms to influence language growth. *Early Childhood Research Quarterly, 53*, 1–10. 10.1016/j.ecresq.2020.02.002.

Crosnoe, R., Purtell, K. M., Davis-Kean, P., Ansari, A., & Benner, A. D. (2016). The selection of children from low-income families into preschool. *Developmental Psychology, 52*(4), 599–612. 10.1037/dev0000101.

(2001). In Dickinson, D. K., & Tabors, P. O. (Eds.), *Beginning literacy with language: Young children learning at home and school* (p. 409). Paul H Brookes Publishing. pp. xx.

Duflo, E., Dupas, P., & Kremer, M. (2011). Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in Kenya. *American Economic Review, 101*(5), 1739–1774. 10.1257/aer.101.5.1739.

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., Pagani, L. S., Feinstein, L., Engel, M., Brooks-Gunn, J., Sexton, H., Duckworth, K., & Japel, C. (2007). School readiness and later achievement. *Developmental Psychology, 43*(6), 1428–1446. 10.1037/0012-1649.43.6.1428.

Duncan, G. J., & Murnane, R. J. (2011). *Whither Opportunity?: Rising Inequality, Schools, and Children's Life Chances*. Russell Sage Foundation.

Early, D. M., Iruka, I. U., Ritchie, S., Barbarin, O. A., Winn, D.-M. C., Crawford, G. M., Frome, P. M., Clifford, R. M., Burchinal, M., & Howes, C. (2010). How do pre-kindergarteners spend their time? Gender, ethnicity, and income as predictors of experiences in pre-kindergarten classrooms. *Early Childhood Research Quarterly, 25*(2), 177–193.

Enders, C. K. (2013). Dealing with missing data in developmental research. *Child Development Perspectives, 7*(1), 27–31. 10.1111/cdep.12008.

Eslava, M., Deaño, M., Alfonso, S., Conde, Á., & García-Señorán, M. (2016). Family context and preschool learning. *Journal of Family Studies, 22*(2), 182–201. 10.1080/13229400.2015.1063445.

Finch, J. E., Garcia, E. B., Sulik, M. J., & Obradović, J. (2019). Peers matter: Links between classmates' and individual students' executive functions in elementary school. *AERA Open, 5*(1), Article 2332858419829438. 10.1177/2332858419829438.

Fletcher, J. (2010). Spillover effects of inclusion of classmates with emotional problems on test scores in early elementary school. *Journal of Policy Analysis and Management, 29*(1), 69–83. 10.1002/pam.20479.

Foster, T. J., Burchinal, M., & Yazejian, N. (2020). The relation between classroom age composition and children's language and behavioral outcomes: Examining peer effects. *Child Development, 91*(6), 2103–2122. 10.1111/cdev.13410.

Fuhs, M. W., Nesbitt, K. T., Farran, D. C., & Dong, N. (2014). Longitudinal associations between executive functioning and academic skills across content areas. *Developmental Psychology, 50*(6), 1698–1709. 10.1037/a0036633.

Fuligni, A. S., Howes, C., Huang, Y., Hong, S. S., & Lara-Cinisomo, S. (2012). Activity settings and daily routines in preschool classrooms: Diverse experiences in early learning settings for low-income children. *Early Childhood Research Quarterly, 27*(2), 198–209.

Gottfried, M. A., & Kirksey, J. J. (2019). General education teachers' math instructional practices in kindergarten classrooms with and without students with emotional and behavioral disabilities. *Teaching and Teacher Education, 77*, 309–320. 10.1016/j.tate.2018.10.019.

Halle, T. G., Hair, E. C., Wandner, L. D., & Chien, N. C. (2012). Profiles of school readiness among four-year-old Head Start children. *Early Childhood Research Quarterly, 27*(4), 613–626. 10.1016/j.ecresq.2012.04.001.

Hanushek, E. A., Kain, J. F., Markman, J. M., & Rivkin, S. G. (2003). Does peer ability affect student achievement? *Journal of Applied Econometrics, 18*(5), 527–544. 10.1002/jae.741.

Henry, G. T., & Rickman, D. K. (2007). Do peers influence children's skill development in preschool? *Economics of Education Review, 26*(1), 100–112. 10.1016/j.econedurev.2005.09.006.

Hong, G., Corter, C., Hong, Y., & Pelletier, J. (2012). Differential effects of literacy instruction time and homogeneous ability grouping in kindergarten classrooms: Who will benefit? Who will suffer? *Educational Evaluation and Policy Analysis, 34*(1), 69–88. 10.3102/0162373711424206.

Hulsey, L. K., Aikens, N., Kopack, A., West, J., Moiduddin, E., & Tarullo, L. (2011). Head Start children, families, and programs: Present and past data from FACES. OPRE Report 2011-33a. *Administration for Children & Families*. Administration for Children & Families https://eric.ed.gov/?id=ED539261.

Justice, L. M., Jiang, H., Purtell, K. M., Lin, T.-J., & Ansari, A. (2021). Academics of the early primary grades: Investigating the alignment of instructional practices from pre-K to Third Grade. *Early Education and Development*, 1–19. 10.1080/10409289.2021.1946762.

Justice, L. M., Logan, J. A. R., Lin, T.-J., & Kaderavek, J. N. (2014). Peer effects in early childhood education: Testing the assumptions of special-education inclusion. *Psychological Science, 25*(9), 1722–1729. 10.1177/0956797614538978.

Justice, L. M., Petscher, Y., Schatschneider, C., & Mashburn, A. (2011). Peer effects in preschool classrooms: Is children's language growth associated with their classmates' skills? *Child Development, 82*(6), 1768–1777. 10.1111/j.1467-8624.2011.01665.x.

Kim, S., Lin, T.-J., Chen, J., Logan, J., Purtell, K. M., & Justice, L. M. (2020). Influence of teachers' grouping strategies on children's peer social experiences in early elementary classrooms. *Frontiers in Psychology, 11*, 3583. 10.3389/fpsyg.2020.587170.

Kohl, K., Bihler, L., Agache, A., Leyendecker, B., & Willard, J. (2021). Do peers matter? Peer effects on young children's vocabulary gains in German classrooms. *Journal of Educational Psychology*. 10.1037/edu0000522.

Kuhn, L. J., Willoughby, M. T., Wilbourn, M. P., Vernon-Feagans, L., & Blair, C. B.The Family Life Project Key Investigators. (2014). Early communicative gestures prospectively predict language development and executive function in early childhood. *Child Development, 85*(5), 1898–1914. 10.1111/cdev.12249.

Kuzmina, Y., & Ivanova, A. (2018). The effects of academic class composition on academic progress in elementary school for students with different levels of initial academic abilities. *Learning and Individual Differences, 64*, 43–53. 10.1016/j.lindif.2018.04.004.

Ledford, J. R., & Wolery, M. (2015). Observational learning of academic and social behaviors during small-group direct instruction. *Exceptional Children, 81*(3), 272–291. 10.1177/0014402914563698.

Lin, T.-J., Justice, L. M., Paul, N., & Mashburn, A. J. (2016). Peer interaction in rural preschool classrooms: Contributions of children's learning-related behaviors, language and literacy skills, and problem behaviors. *Early Childhood Research Quarterly, 37*, 106–117. 10.1016/j.ecresq.2016.04.001.

Luyten, H., Schildkamp, K., & Folmer, E. (2009). Cognitive development in Dutch primary education, the impact of individual background and classroom composition. *Educational Research and Evaluation, 15*(3), 265–283. 10.1080/13803610902955333.

Magnuson, K., & Waldfogel, J. (2016). Trends in income-related gaps in enrollment in early childhood education: 1968 to 2013. *AERA Open*: 2 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6110399/.

Mashburn, A. J., Justice, L. M., Downer, J. T., & Pianta, R. C. (2009). Peer effects on children's language achievement during Pre-Kindergarten. *Child Development, 80*(3), 686–702. 10.1111/j.1467-8624.2009.01291.x.

McGrew, K. S., & Woodcock, R. W. (1989). *Technical manual. Woodcock-Johnson PsychoEducational Battery–Revised*. Rolling Meadows, IL: Riverside Publishing.

McWayne, C. M., Fantuzzo, J. W., & McDermott, P. A. (2004). Preschool competency in context: An investigation of the unique contribution of child competencies to early academic success. *Developmental Psychology, 40*(4), 633–645. 10.1037/0012-1649.40.4.633.

Montroy, J. J., Bowles, R. P., Skibbe, L. E., McClelland, M. M., & Morrison, F. J. (2016). The development of self-regulation across early childhood. *Developmental Psychology, 52*(11), 1744–1762. 10.1037/dev0000159.

National Center for Education Statistics. (2018). *Digest of Education Statistics, 2019*. National Center for Education Statistics https://nces.ed.gov/programs/digest/d19/tables/dt19_202.20.asp.

Nelson, T. D., Nelson, J. M., James, T. D., Clark, C. A. C., Kidwell, K. M., & Espy, K. A. (2017). Executive control goes to school: Implications of preschool executive performance for observed elementary classroom learning engagement. *Developmental Psychology, 53*(5), 836–844. 10.1037/dev0000296.

Nesbitt, K. T., Farran, D. C., & Fuhs, M. W. (2015). Executive function skills and academic achievement gains in prekindergarten: Contributions of learning-related behaviors. *Developmental Psychology, 51*(7), 865–878. 10.1037/dev0000021.

Pianta, R., Howes, C., Burchinal, M., Bryant, D., Clifford, R., Early, D., & Barbarin, O. (2005). Features of pre-Kindergarten programs, classrooms, and teachers: Do they predict observed classroom quality and child-teacher interactions? *Applied Developmental Science, 9*(3), 144–159. 10.1207/s1532480xads0903_2.

Pianta, R. C., Whittaker, J. E., Vitiello, V., Ansari, A., & Ruzek, E. (2018). Classroom process and practices in public pre-K programs: Describing and predicting educational opportunities in the early learning sector. *Early Education and Development, 29*(6), 797–813. 10.1080/10409289.2018.1483158.

Purpura, D. J., Hume, L. E., Sims, D. M., & Lonigan, C. J. (2011). Early literacy and early numeracy: The value of including early literacy skills in the prediction of numeracy development. *Journal of Experimental Child Psychology, 110*(4), 647–658. 10.1016/j.jecp.2011.07.004.

Reynolds, A. J., Temple, J. A., Ou, S.-R., Arteaga, I. A., & White, B. A. B. (2011). School-based early childhood education and age-28 well-being: Effects by timing, dosage, and subgroups. *Science, 333*(6040), 360–364. 10.1126/science.1203618.

Ribeiro, L. A., Zachrisson, H. D., & Dearing, E. (2017). Peer effects on the development of language skills in Norwegian childcare centers. *Early Childhood Research Quarterly, 41*, 1–12. 10.1016/j.ecresq.2017.05.003.

Rubin, K., Bukowski, W., & Parker, J. (2006). Handbook of Child Psychology. In *Handbook of Child Psychology* (pp. 571–645). Vol. 3. 10.1002/9780470147658.chpsy0310.

Rjosk, C., Richter, D., Hochweber, J., Lüdtke, O., Klieme, E., & Stanat, P. (2014). Socioeconomic and language minority classroom composition and individual reading achievement: The mediating role of instructional quality. *Learning and Instruction, 32*, 63–72. 10.1016/j.learninstruc.2014.01.007.

Sawyer, B., Atkins-Burnett, S., Sandilos, L., Hammer, C. S., Lopez, L., & Blair, C. (2018). Variations in classroom language environments of preschool children who are low income and linguistically diverse. *Early Education and Development, 29*(3), 398–416. 10.1080/10409289.2017.1408373.

Schmerse, D. (2021). Peer effects on early language development in dual language learners. *Child Development, 92*(5), 2153–2169. 10.1111/cdev.13588.

Schmitt, S. A., Geldhof, G. J., Purpura, D. J., Duncan, R., & McClelland, M. M. (2017). Examining the relations between executive function, math, and literacy during the transition to kindergarten: A multi-analytic approach. *Journal of Educational Psychology, 109*(8), 1120–1140. 10.1037/edu0000193.

Shen, T., & Konstantopoulos, S. (2022). Incorporating complex sampling weights in multilevel analyses of educational data. *Practical Assessment, Research, and Evaluation, 27*(13), 1–22. 10.7275/PTKW-5816.

Sills, J., Rowse, G., & Emerson, L.-M. (2016). The role of collaboration in the cognitive development of young children: A systematic review. *Child: Care, Health & Development, 42*(3), 313–324. 10.1111/cch.12330.

Skibbe, L. E., Gerde, H. K., Wright, T. S., & Samples-Steele, C. R. (2016). A content analysis of phonological awareness and phonics in commonly used Head Start curricula. *Early Childhood Education Journal, 44*(3), 225–233. 10.1007/s10643-015-0703-8.

Skibbe, L. E., Phillips, B. M., Day, S. L., Brophy-Herb, H. E., & Connor, C. M. (2012). Children's early literacy growth in relation to classmates' self-regulation. *Journal of Educational Psychology, 104*(3), 541–553. 10.1037/a0029153.

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*(2), 173–187. 10.1016/j.ecresq.2007.01.002.

Tomlinson, C. A., Brighton, C., Hertberg, H., Callahan, C. M., Moon, T. R., Brimijoin, K., Conover, L. A., & Reynolds, T. (2003). Differentiating instruction in response to student readiness, interest, and learning profile in academically diverse classrooms: A review of literature. *Journal for the Education of the Gifted, 27*(2–3), 119–145. 10.1177/016235320302700203.

Vallotton, C., & Ayoub, C. (2011). Use your words: The role of language in the development of toddlers' self-regulation. *Early Childhood Research Quarterly, 26*(2), 169–181. 10.1016/j.ecresq.2010.09.002.

Vukovic, R. K., & Lesaux, N. K. (2013). The language of mathematics: Investigating the ways language counts for children's mathematical development. *Journal of Experimental Child Psychology, 115*(2), 227–244. 10.1016/j.jecp.2013.02.002.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Weiland, C., & Yoshikawa, H. (2014). Does higher peer socio-economic status predict children's language and executive function skills gains in prekindergarten? *Journal of Applied Developmental Psychology, 35*(5), 422–432. 10.1016/j.appdev.2014.07.001.