Submitted to *Journal of Applied Developmental Psychology*, November 13, 2017
Revision submitted March 26, 2018

CONTINUITY AND CHANGES IN CLASSROOM AGE COMPOSITION AND ACHIEVEMENT IN HEAD START

Arya Ansari
University of Virginia

Kelly M. Purtell
The Ohio State University

Running Head: Classroom Age Composition
Word Count: 147 (abstract), 7019 (text and references)
4 Tables, 0 Figures

Ansari, A., & Purtell, K. M. (in press). Continuity and changes in classroom age composition and achievement in Head Start. *Journal of Applied Developmental Psychology.* Advance online publication. doi: 10.1016/j.appdev.2018.07.002

* Correspondence concerning this article should be addressed to the first author at the Center for Advanced Study of Teaching and Learning, University of Virginia, PO Box 800784, Charlottesville, VA 22908-0784 (email: aa2zz@eservices.virginia.edu). This research was supported by a grant from the American Educational Research Association which receives funds for its "AERA Grants Program" from the National Science Foundation under NSF Grant #DRL-0941014 and the Institute of Education Sciences, U.S. Department of Education (R305B130013, University of Virginia). Opinions reflect those of the authors and do not necessarily reflect those of the granting agencies.

Abstract

Using data from the Family and Child Experiences Survey 2009 Cohort (*n* = 1,073), this study considered the implications of mixed-age education for young children's academic achievement when they experienced continuity and/or changes in classroom age composition across two years in Head Start (at age 3 and age 4). Results from these analyses revealed that children in classrooms with a greater number of younger children during their second year in Head Start exhibited fewer gains in mathematics and language and literacy. Additionally, children who transitioned from being in classrooms with largely older classmates during year one to classrooms with largely same-age peers during year two exhibited greater gains in academics than children who experienced two years of mixed-age classrooms. Stability in children's teachers, one of the hallmarks of mixed-age programming, was not associated with children's academic achievement nor did it attenuate the negative consequences of mixed-age classrooms.

Keywords: classroom age composition; Head Start; FACES 2009; academic achievement

Continuity and Changes in Classroom Age Composition and Achievement in Head Start

The growing recognition of early childhood as a critical developmental period, one that has lasting influences throughout the life course, has spearheaded the expansion of preschool education for both 3- and 4-year-olds across the country (Duncan & Magnuson, 2014; Yoshikawa et al., 2013). One way in which programs can expand is by including children of both ages in the same classroom (e.g., 3- and 4-year-olds). Indeed, recent national estimates reveal that over three quarters of children in Head Start—the largest federally funded preschool program in the United States—are enrolled in mixed-age classrooms (Ansari et al., 2016). Other national estimates also reveal that the majority of public and private preschool programs that serve 3- and/or 4-year-olds have over a 12-month difference in age between the oldest and youngest student in the classroom (National Survey of Early Care and Education, 2012, authors' calculations), suggesting that a large proportion of preschoolers in the United States experience mixed-age education. Even with the large number of mixed-age classrooms across the country, it remains unclear whether these programs are designed to provide children with multiple years of developmentally appropriate educational opportunities. In fact, to date, much of the mixed-age education literature has focused on children's classroom experiences during their first year in the program and, as a result, what happens to children who go on to experience a second year in the these types of programs, when they are often the older children, is unclear.

Some scholars also argue that continuity in children's peers and caregivers across school years—one of the hallmarks of mixed-age classrooms—is beneficial for children's early learning and development (Katz, Evangelou, & Hartmann , 1990; Veenman, 1995), whereas others contend that these type of settings are likely less conducive for children given the demands associated with mixed-age education (Mason & Burns, 1996). The purpose of the current

investigation is to address these competing hypotheses by examining the academic implications of having different-age peers across two school years and the role of teacher continuity in evaluating the benefits and drawbacks of mixed-age classrooms in Head Start. This type of longitudinal empirical inquiry is of growing importance given the fact that some scholars and professional and national associations have been promoting mixed-age groupings as beneficial for children in early childhood programs even when there has been little evidence either in support of (or against) this type of educational programming (Katz et al., 1990).

Theoretical Underpinnings of Mixed-age Education and its Implications

Our interest in understanding the unique influence of classroom age composition across school years is grounded in several long-standing educational and developmental models. The overarching framework for our work is based on Bronfenbrenner’s bioecological theory, which emphasizes the importance of contextual influences on children’s early learning and development (Bronfenbrenner & Morris, 2006). In line with this theory, the current investigation considers how specific proximal processes within Head Start classrooms—age composition and teacher continuity—can influence children’s academic achievement. Bandura’s (1986) social learning theory and Vygotsky’s (1978) theory of cognitive development also shape our developmental framework for this study as they both contend that one of the primary mechanisms through which development occurs in early childhood programs is through interactions between children and their classmates. Younger children can observe older and more skilled children in the classroom and mimic their behaviors and actions and older children can scaffold younger children who, in turn, can cement their own skills and knowledge. When taken together, these transactional processes among peers can constitute one of the key mechanisms through which early childhood programs impact children’s academic and social-behavioral development (e.g., Henry & Rickman, 2007; Justice, Logan, Lin, & Kaderavek, 2014; Mashburn,

Justice, Downer, & Pianta, 2009; Ribeiro, Zachrisson, & Dearing, 2018).

Despite its plausible theoretical underpinnings and endorsements by national organizations, the evidence behind these types of classrooms has been largely inconclusive when looking at children's early academic and socioemotional development, with some early childhood scholars documenting positive impacts (Blasco, Bailey & Burchinal, 1993; Goldman, 1981; Guo, Tompkins, Justice, & Petscher, 2014; Justice, Logan, Purtell, Bleses, & Hojen, 2018) and others documenting null or negative associations (Ansari et al., 2016; Bell et al., 2013; Moller, Forbes-Jones, & Hightower, 2008; Urberg & Kaplan, 1986; Winsler, Caverly, Willson-Quayle, Carlton, Howell, & Long , 2002). Studies of the academic benefits of mixed-age (or multi-grade) classrooms serving children in the elementary grades has also been largely ambiguous, with some studies documenting academic effects that were positive, negative, and statistically indistinguishable (Ansari, 2017; Pratt, 1986; Proehl, Douglas, Elias, Johnson, & Westsmith, 2013; Thomas, 2012; Veenman, 1995; Way, 1981). Thus, across the educational spectrum, the empirical support for mixed-age education has been largely inconclusive.

Even with the conflicting empirical evidence in the existing literature, a study by Ansari et al. (2016) is of note as it represents the first national study of mixed-age classrooms in the United States for first time Head Start attendees. More specifically, Ansari et al.f (2016) documented sizeable negative associations between mixed-age classrooms for newly enrolled 4-year-olds' math and language and literacy learning, and found that classroom age composition did *not* have mean-level associations with the early academic success of 3-year-olds. Practically speaking, the drawbacks of mixed-age classrooms for first-time 4-year-old Head Start attendees amounted to approximately four to five months of academic development when they attended classrooms that enrolled an equal number of 3- and 4-year-olds. There were, however, no benefits or drawbacks of mixed-age classrooms for children's socioemotional development

(Ansari et al.,, 2016).

To date, however, no studies, including Ansari et al. (2016), have examined what happens to the 3-year-olds in their second year of Head Start. That is, what happens to children who spend two years in mixed-age classrooms? This type of empirical analysis has important implications for policy and practice as the majority of children who attend Head Start at age 3 remain in the program for a second year as 4-year-olds (Puma et al., 2010). In fact, this empirical inquiry into mixed-age education can point to one of the potential reasons why children who experience preschool—especially Head Start—at ages 3 and 4 make greater gains during their first year than in their second (Jenkins, Farkas, Duncan, Burchinal, & Vandell , 2016; Yoshikawa et al., 2013). Accordingly, our first research objective was to assess the academic implications of mixed-age classrooms for 3-year-olds during their second year in the Head Start program. Based on the work of Ansari et al. (2016), we expected that children who were enrolled in classrooms with a greater share of younger classmates during their second year in Head Start as 4-year-olds would demonstrate smaller gains in areas of literacy and math.

Continuity and Changes in Classroom Age Composition

As part of the current investigation, we also consider the implications of mixed-age education for children's academic achievement when they experience continuity and/or changes in classroom age composition across two years in the Head Start program. In other words, what happens to children in the program who transition from classrooms where their classmates are largely the same age during year one to classrooms where there are largely different age peers in year two? Alternatively, what happens to children who are enrolled in classrooms with largely different age peers across both school years? Taking a multi-year perspective on children's classroom experiences is grounded in developmental theory (Bronfenbrenner & Morris, 2006) and can allow for a more nuanced understanding of the different ways in which classroom

dynamics shape children's academic success.

Perhaps mixed-age classrooms require children to be present for multiple years to reap the maximum benefit (Lillard, 2016). For example, being a younger child in a classroom in one year—although not benefiting them immediately (Ansari et al., 2016)—may motivate children to be like one of their older classmates (Winsler et al., 2002) and, thus, in their second year, these children may start school with skills that allow them to more effectively scaffold for their younger peers and cement their own skills and knowledge. On the other hand, a recent report from the Department of Education clearly shows that the vast majority of educators feel underprepared to individualize and differentiate their instruction (Manship, Farber, Smith, & Drummond, 2016), and therefore, it is also conceivable that regardless of children's prior experiences in mixed-age settings, these types of environments are not optimal for their early learning. Regardless of the outcome, this type of longitudinal empirical inquiry is of utmost importance because these transitions across years likely represent qualitatively different experiences that might alter the meaning of mixed-age education for children in any given year. Examining 3-year-olds in the Head Start program in particular presents a unique opportunity to assess the implications of such transitions and we address this objective concerning children's classroom transitions in two different ways.

First, we use continuous measures of classroom age composition at years one and two, and then in the second set of analyses, we cross two sets of categorical indicators of high and low levels of same-age peers (for similar methods see: Burchinal, Vandell, & Belsky, 2014). The first set of analyses captures whether each unit increase in different age peers during year one moderates the effect of classroom age composition during year two (i.e., a linear effect). In the second set of analyses we test for the multiplicative effects of high (versus low) levels of different age peers across the two Head Start years using standard deviation cut points (see also,

Ansari et al., 2016). In doing so, the second set of analyses captures potential non-linear effects that tap into qualitatively different classroom transitions (e.g., transitioning from different age classrooms at age 3 to same-age classrooms at age 4).

As discussed by Weiland and Yoshikawa (2014), there is no consensus for selecting a threshold, but possibilities include inflection points, conceptually defined points, empirically identified points, in addition to nonlinear methods. For the purposes of the present investigation, we test conceptually defined points that correspond to classrooms in which >30% of children are of a different age. These estimates were based on prior work with these data that suggest that classrooms where 20-30% of children are of a different age represent qualitatively different experiences (Ansari et al., 2016). It is important to note that similar thresholds of 25-30% have also been used to demarcate preschool classrooms with economic diversity (Miller, Votruba-Drzal, McQuiggan, Shaw, 2017; Weiland & Yoshikawa, 2014). However, given the largely exploratory nature of this objective and the limited work in this area, we also examined two other thresholds that corresponded to classrooms in which: (a) >50% of children's classmates and (b) >70% of children's classmates were of a different age. And given the exploratory nature of this question, we did not make directional hypotheses regarding the implications of children's classroom transitions for their math and reading achievement.

Continuity of Caregivers and Teachers in Mixed-age Classrooms

Finally, one of the hallmarks of the mixed-age model of education is the continuity of teachers and caregivers, which allows children to develop stronger relationships with their teachers (McMullen, Yun, Mihai, & Kim, 2016; Veenman, 1995). The practice of teacher continuity is grounded in attachment theory (Ainsworth, Blehar, Waters, & Wall, 1978) and scholars have found that building stronger relationships with teachers is linked with more optimal school readiness gains (Elicker & Fortner-Wood, 1995; Tran & Winsler, 2011). Similar

to mixed-age environments more generally, continuity of teachers and caregivers has received a number of endorsements by several national organizations (e.g., the National Association for the Education of Young Children; the National Head Start Association and Early Head Start; Zero to Three) despite the limited research in this area (Sosinsky, Ruprect, Horm, Kriener-Althen, & Halle, 2016). In fact, proponents of mixed-age education have also long contended that one of the primary reasons why children in mixed-age classrooms are more likely to succeed is their exposure to continuity in caregivers (Katz et al., 1990; Veenman, 1995).

To our knowledge, however, there has been limited empirical inquiry into the potential benefits of teacher continuity across school years within the context of mixed-age classrooms, especially in large-scale and routinely implemented programs. That is, although there is evidence that teacher-turnover *within* a year results in less optimal school performance for young children (e.g., Markowitz & Bassok, 2018; Tran & Winsler, 2013), whether teacher continuity *across* school years facilitates children's early learning and development in the context of mixed-age classrooms is unknown. In the few known studies that have examined continuity of teachers across school years, the evidence has been mixed, with some early childhood scholars finding benefits of teacher continuity for children's socioemotional development (e.g., Howes & Hamilton, 1993; Howes et al.,1999; Owen, Klausli, Mata-Otero, O'Brien-Caughy, 2008; Ruprecht, Elicker, & Choi, 2016) and others documenting mixed effects (Horm et al., 2018).

In terms of young children's academic achievement in particular, continuity in caregivers across school years might promote their academic learning and buffer the harmful effects of mixed-age classrooms because this practice allows teachers to have more familiarity with children's abilities and, therefore, these teachers might be more successful in matching the content of their curriculum with children's individual needs. However, in one of the only known studies in this area, Horm and colleagues (2018) found no evidence to suggest that continuity of

caregivers and teachers across years was linked with improvements in young children's language and literacy development.

Given the dearth of empirical inquiry in this area, especially in the context of mixed-age classrooms, our final research objective was to determine whether: (a) teacher stability across two years of preschool is associated with children's early academic learning; and (b) whether continuity in teachers attenuates (or amplifies) the effects of mixed-age classrooms. In light of some of the existing educational literature that suggests that continuity in caregivers within (and across) years is potentially positive for children's early learning and adjustment (e.g., Howes & Hamilton, 1993; Howes et al.,1999; Markowitz & Bassok, 2018; Phillips et al., 1994; Tran & Winsler, 2011), we hypothesized that having the same teacher across two years of Head Start would mitigate some of the negative effects of mixed-age classrooms as this would allow teachers to better align the classroom environment to children's individual needs.

The Current Study

When taken together, the current investigation builds on prior work in this area and takes a number of steps to advance our knowledge of mixed-age education, which is the dominant model of educating young children in many early childhood programs, including in Head Start (Ansari et al., 2016; National Survey of Early Care and Education, 2012). Specifically, we first describe: (a) the frequency of mixed-age classroom participation across two years of Head Start; (b) the extent to which children in Head Start experience an increase or decrease in the number of same- and different-age peers across school years; and (c) the number of Head Start attendees who remain with the same teacher for two years (Aim 1). We then examine children's math and reading achievement as a function of mixed-age classrooms across two school years (Aim 2) with a focus on continuity and changes in their classroom experiences (Aim 3). In addressing these objectives, this investigation also considers the role of teacher continuity across school

years within mixed-age classrooms in Head Start (Aim 4).

Given the large number of mixed-age classrooms in early childhood programs across the United States (Ansari et al., 2016; National Survey of Early Care and Education, 2012), the results of the current investigation can have important implications for the implementation of mixed-age education. These findings can also provide important insight into why prior evaluations of Head Start have found that older children do not benefit as much from their enrollment in the program as do younger children (Puma et al., 2010) and why the added academic benefits of each additional year of Head Start tends to diminish (Jenkins et al., 2016; Yoshikawa et al., 2013).

Method

The Family and Child Experiences Survey (FACES) 2009 investigation is a nationally representative sample of 3,349 3- and 4-year-old first time Head Start attendees across 60 programs and 486 classrooms. Children and families who participated in the FACES study were sampled from all 50 states and the District of Columbia and assed either annually or biannually from entry into Head Start through the end of kindergarten (see Moiduddin et al., 2012 for sampling information). For the purposes of the current study, we focused on the 3-year-old cohort (n = 1,954) and their experiences across the 2009-2010 school year (i.e., children's 3 year old year in Head Start) and 2010-2011 school year (i.e., children's 4 year old year in Head Start), as these periods capture the two years children were enrolled in the Head Start program.

We restricted our sample to the 3 year olds—defined as those children who were two years away from kindergarten entry and participated in the data collection through the spring of 2011. These criteria resulted in 823 children being dropped. Even though we restricted our sample to children who were eligible for kindergarten two years after their entry into Head Start, 15 children enrolled in kindergarten anyway and, thus, were dropped. Finally, we required that

children had not switched classrooms during year one, resulting in an additional 43 children being dropped from our analytic sample. When taken together, these exclusion criteria resulted in a final sample of 1,073 children who were drawn from 402 classrooms across 118 centers. The majority of children in our final sample were identified as either Latino (37%) or Black (34%). During the spring of 2011, children were, on average, 5.01 years of age (*SD* = 0.31). Roughly 54% of these children had mothers who were unemployed and 35% of children had mothers with less than a high school education (for other sample demographics, see Table 1).

It is important to note that our final analytic sample of 1,073 children was similar to the 1,644 children who only participated in Head Start for one year (see: Ansari et al., 2016). However, even when weighted to account for cross-wave attrition, children who were excluded because they did not have a valid weight in Year 1 were more likely to be White, whereas children who participated throughout Year 1 and Year 2 were more likely to be Latino.

Measures

Weighted descriptives for all focal variables are provided in Tables 1 and 2.

Classroom age composition during year one. During the fall of 2009, teachers reported how many children were in their classroom, and how many were 3, 4, and 5 years of age. Because there were only a small number of 5-year-olds at the beginning of the school year, we dichotomized children as 3 years old or younger or 4 years or older (see also, Ansari et al., 2016; Moiduddin et al., 2012). Then, we then divided the number of 3-year-olds by the overall class size to create our indicator of classroom age composition during year one.

Classroom age composition during year two. Unlike the year one reports of classroom age composition, during the second Head Start year we had access to teacher reports of classroom age composition at the end of the year (spring of 2011). Specifically, teachers reported how many children were in their classroom and how many children were 3 years of age or

younger, 4 years of age, and 5 years of age. We divided the number of children across each age group by the class size to create our focal independent variables of classroom age composition during year two (i.e., the percent of 3-, 4-, and 5-year-old children).

Teacher stability. Teacher stability across the two Head Start years was coded based on the unique teacher identification numbers. Teacher change was defined as the children who had two different teacher codes across these two years (0= *child experienced a teacher change*), whereas teacher stability was defined as the children who had the same teacher identification number across the two time points (1= *child had the same teacher*).

Children's academic achievement. At each wave of data collection, children's language and literacy skills were directly assessed with the Woodcock-Johnson Letter Word Identification (α's = 0.85-0.93) and Spelling tests (α's = 0.79-0.87; Woodcock, McGrew, & Mather, 2001) and the Peabody Picture Vocabulary Test (α's = 0.93-0.97; Dunn & Dunn, 1997). Similar to prior studies (Ansari et al., 2016; Duncan et al., 2015; Jenkins et al., 2018), we created a composite for early reading achievement because each of the outcomes followed a similar pattern. Children's math skills were also directly assessed with the Woodcock-Johnson Applied Problems subscale (Woodcock et al., 2001) and with assessments from the Early Childhood Longitudinal Study-Birth Cohort (α's = .80-0.92; Snow et al., 2007). Children who came from non-English-speaking households were screened for their English proficiency prior to each assessment, and if they did not pass the test, they were assessed in Spanish (2-26% across waves).

Covariates. All of the models discussed below control for an extensive set of child and family covariates that are listed in Table 1. The covariates in our models are largely informed by biecological theory (Bronfenbrenner & Morris, 2006) and tap into aspects of children's lives that prior studies have shown are linked with their development and may be associated with the types of classrooms that they experience. Thus, in controlling for these covariates, our statistical

models isolate the associations between the age composition of classrooms from other factors that might covary with both children's early academic learning and the classrooms that serve different age children. These child and family covariates were almost always reported on by children's parents and include aspects of *household structure* (i.e., mothers' marital status and household size), *socioeconomic status and parents' mental health* (i.e. mothers' education, ratio of income to poverty, employment, and depressive symptoms [as measured by the Center for Epidemiologic Studies Depression Scale; Radloff, 1977]), and *child characteristics* (i.e., gender, race/ethnicity, age at Head Start entry, home language, and language of assessment).

As part of these covariates, we also included classroom-level factors from both Head Start years in order to isolate the unique "effects" of classroom age composition (for descriptives, see Table 2). More specifically, in our models we also control for: child/teacher ratios, child/adult ratios, class size, teachers' depressive symptoms (as measured by the Center for Epidemiologic Studies Depression Scale; Radloff, 1977), hours of school per week, multilingual instruction (0 = *no*, 1 = *yes*), teachers' years of experience, teachers' years of education, whether teachers' degrees are in early childhood education, teachers' race/ethnicity, teachers' hourly salary, teachers' benefits (e.g., paid leave, sick leave), an indicator for whether teachers taught in multiple classrooms (0 = *no*, 1 = *yes*), and teacher-child interaction quality as measured by the Classroom Assessment Scoring System (Pianta, La Paro, & Hamre, 2008). Teachers generally reported on these classroom covariates.

Finally, all of our statistical models presented below adjusted for children's academic skills during the fall of 2009 and spring of 2010, which is recognized as one of the strongest adjustments for omitted variable bias (for more details see: National Institute of Child Health and Human Development Early Child Care Research Network & Duncan, 2003). We also controlled for the length of months between assessments.

Analytic Strategy

All analyses were estimated with Stata (StataCorp, 2009) and Mplus (Muthén & Muthén, 1998-2013). To address *Aim 1* and provide a snapshot of the focal variables of interest, we estimated weighted descriptive statistics (e.g., means and standard deviations) in Stata. Then, to address our focal research objectives, we estimated five sequential models within an OLS regression framework. Specifically, to address *Aim 2* of this investigation (i.e., assess the main effects of classroom age composition during year two) we estimate a series of models in the Mplus program. Our first model compares the experiences of children when they were enrolled in Head Start classrooms with different proportions of younger and older children. As part of this effort, we also tease apart the proportion of younger children who were 3 and 4 years of age during the second Head Start year.

In order to address *Aim 3* (i.e., evaluate the implications of continuity and changes in classroom age composition), our next set of models included interaction terms between the continuous indicators of classroom age composition during years one and two. We also estimated models that considered the transitions children experienced across school years using standard deviation cut points as a means of looking at qualitatively different classroom transitions. Our final set of analyses included interaction terms between classroom age composition during year two and teacher stability as a means of addressing *Aim 4* (i.e., consider the implications of teacher stability across school years within mixed-age classrooms in Head Start).

In addition to the details outlined above, all models: (a) were clustered at the classroom level using TYPE = COMPLEX to account for non-independence in child outcomes (ICCs = .20-.25); (b) were weighted to account for variation in the probability of selection as well as eligibility and attrition (weight = PRA13WT); (c) employed full information maximum likelihood estimation (Schafer & Graham, 2002) to address missing data (the average amount of

missing data was 9% and missing data for any one variable ranged from 0 to roughly 45%); and (d) controlled for all of the covariates listed in Tables 1 and 2. Finally, it is important to note that all continuous variables have been standardized to have a mean of zero and standard deviation of one and, thus, all coefficients correspond to effect sizes (e.g., a Cohen's *d*).

Results

Nature and variability of classroom age composition and teacher continuity. At the beginning of Year 1 of Head Start, 58% of children's classmates were 3 years of age and 42% of children's peers were 4 years of age. At the end Year 2 when children were, on average, 5 years of age, 14% of children's peers were 3 years of age, 46% were 4 years of age, and 39% were 5 years of age. Roughly four in ten children also had the same teacher across both years in Head Start. For a breakdown of the classroom transitions based on the three thresholds, see Table 2.

Implications of age composition in year two of Head Start. As can be seen in Models 1 of Table 3, we found that when there was a greater share of younger children, the older children in the classrooms displayed fewer gains in reading and math with effect sizes corresponding to 7% and 5% of a standard deviation, respectively. When we separated the proportion of younger children into two variables that captured the proportion of children who were 3 years of age and the proportion who were 4 years of age, we found that the associations were negative for both groups as compared with 5-year-olds, but only the contrast with 3-year-olds reached conventional levels of statistical significance (and only for reading; see Models 2). These coefficients were *not* significantly different from one another (Wald x^2 = |0.07-1.41|, *ns*). In other words, the negative implications of mixed-age education were driven equally by the both groups of younger children in the classroom, supporting our use of a combined indicator.

To put these findings in context, we translate the statistical representation of these effect sizes into a more readily interpretable form of months of development (these estimates are

calculated by dividing the standardized difference in academic test scores by the regression slope of children's age; Bradbury, Corak, Waldfogel, & Washbrook, 2011). On average, children in the study sample demonstrate a 5-7% of a standard deviation increase in reading and math per month and, thus, the negative associations between classroom age composition and children's achievement translate to roughly 0.70 and 1.40 months of lost opportunities in math and reading.

Continuity and changes in age composition. Having established the negative associations between classroom age composition and children's math and language and literacy learning during the second Head Start year, we next examined whether the age composition during their first year also had implications for children's early academic learning a year later. Results from these analyses revealed that there was *no* relation between classroom age composition from the prior school year (Fall 2009-Spring 2010) with changes in children's academic achievement through the spring of the following year (see Models 1 and 2). As part of these analyses, we also examined whether these potential associations were moderated by classroom age composition in year two and found no evidence for moderation (see Model 3).

Although the continuous cross-year classroom composition interactions were not statistically significant, we did find some evidence to suggest that certain transitions based on our thresholds were more beneficial than others (see Table 4). Specifically, children who transitioned from different age classrooms—where they were the younger children in year one—to classrooms that consisted of largely same-age peers in year two, made the stronger gains in reading and math than children who remained in different-age classrooms across both years. Effect sizes ranged from 15-38% of a standard deviation (depending on the threshold specification). And although less consistent, there was also some trend-level evidence to suggest that children who transitioned from same age classrooms in year one to different age classrooms in year two demonstrated weaker reading and math skills than children who transitioned from

different age to same age classrooms, with effect sizes of 13-26% of a standard deviation.

Continuity and changes in children's teachers. Our final set of analyses considered whether having the same teacher across both years in Head Start could buffer the negative associations between classroom age composition and children's academic growth. We found that having the same teacher across both school years did not influence children's academic achievement (see Models 1 and 2 of Table 3) and there was *no* evidence for moderation (see Model 4 of Table 3). Thus, having larger proportions of younger classmates when children were 4 years of age was negatively related to children's academic school performance regardless of children's experiences with stability and/or changes in their teachers.

Discussion

With the continued expansion of early care and education programs serving young children across the country (Duncan & Magnuson, 2013; Phillips et al., 2017; Yoshikawa et al., 2013), there has been a growing need to understand how different classroom factors are linked with children's early learning and development. The current study sought to address important gaps in the extant literature with regards to the potential role of classroom age composition and the early academic achievement of children over the course of two years in the Head Start program. When taken together, the results of the current investigation have three take home messages.

First, despite its theoretical underpinnings (Bandura, 1986; Vygotsky, 1978) and endorsements by national organizations (Katz et al., 1990), the results from the current investigation reveal that mixed-age classrooms may be problematic for the older children in these classrooms. Children who were enrolled in classrooms with a larger share of younger classmates during their second year in Head Start demonstrated fewer gains in reading and math. Although the effect sizes for the associations between classroom age composition and children's academic learning might appear small when compared with conventional standards, these

associations translate to roughly one month of development and are comparable to prior analyses of classroom age composition for first time Head Start attendees (7% of a standard deviation; Ansari et al., 2016). Moreover, the effect sizes of classroom age composition were comparable to the effects of classroom quality documented in prior meta-analyses (Keys et al., 2013), suggesting that classroom age composition has non-negligible effects on children's academic learning as compared with other dimensions of the classroom that have received extensive attention.

Thus, this study both extends and replicates the findings of Ansari et al. (2016) who found that newly enrolled 4-year-olds exhibited less optimal academic achievement in mixed-aged classrooms, by illustrating the implications of mixed-age education for the academic achievement of a different sample of Head Start attendees who experienced a second year in the program. Put a different way, despite having qualitatively different educational experiences, both newly enrolled 4-year-olds and 4-year-olds who had experienced a prior year in the Head Start program demonstrate fewer gains in areas of math and reading when they had a larger number of younger classmates. With the growing emphasis on replication as a key component of scientific inquiry (Duncan, Engel, Claessens, & Dowsett, 2014), our findings serve as an important confirmation of the potential negative effects of mixed-age classrooms for the academic achievement of 4-year-olds in Head Start, regardless of their prior educational experiences.

In light of the associations documented in this study, one might wonder why mixed-age groupings are not beneficial for older children. These negative associations likely have to do with the fact that early childhood educators are generally not well prepared to teach in these types of classrooms and, consequently, they struggle with managing the needs of their students (Manship et al., 2016; Mason & Burns, 1996). That is, teachers in these classrooms are faced with the challenge of adjusting their instructional practices to a wide range of ages and skill

levels relative to classrooms with more uniform age levels, and what may end up happening is that they target their classroom instruction to children at the lower end of the age distribution. Reflecting these possibilities, a study by Ansari and Pianta (2018a) with a separate dataset of pre-K classrooms across eight states found that preschool teachers who taught in classrooms with greater age diversity demonstrated a decrease in instructional, emotional, and organizational support across the school year. Other studies have also found that teachers in mixed-age classrooms spend significantly less time in academic and teacher-directed instruction (Ansari, 2017). Thus, mixed- and single-age classroom environments are likely to be different in important ways that have ramifications for children's academic learning. Even with these possibilities, the specific reasons why older children do less well academically in mixed-age classrooms is not well understood, which is quite surprising given the wide prevalence of preschool classrooms in the United States that serve different age children. For these reasons, future studies should more carefully consider the underlying mechanisms for these associations.

Second, our transition analyses revealed even though children were performing less well academically when they were enrolled in classrooms with a larger share of younger children, the children who transitioned from classrooms where they were the younger children at age 3 (classmates were largely 4- and 5-year-olds) to classrooms where there were largely with same-age peers at age 4 (largely 4- and 5-year-olds) made greater gains in academics as compared with children who remained in different age classrooms. Whereas these associations were observed when looking at thresholds, we found no evidence of moderation when using continuous variables. Although some researchers have advised that continuous variables should not be categorized when testing for interactions (DeCoster, Iselin, & Gallucci, 2009), this approach assumes that the effect of one variable increases linearly with the other. As demonstrated here, the categorical approach, but not the continuous approach, detected significant differences.

To our knowledge, however, this is the first study of the academic development of young children enrolled in mixed-age classrooms across multiple school years. Accordingly, even though Ansari et al. (2016) found that 3-year-olds did not benefit academically from having older classmates during their initial year in Head Start, our results indicate that there may be academic benefits that emerge later on, which are supported (i.e., transitioning to same-age classrooms) and potentially undermined (i.e., remaining in different age classrooms) by children's experiences during the following year. Considering that this study—to our knowledge—is one of the first attempts to understand these classroom transitions, continued work is necessary to determine whether our results replicate using different methodologies in addition to different samples of children and families in different types of early childhood programs. This type of empirical inquiry is all the more important given both the null and significant patterns documented when examining the classroom transitions.

Finally, although the continuity of teachers and caregivers has long been recognized as one of the cornerstones of mixed-age programs (Veenman, 1995), there has been little to no inquiry into the potential benefits of these practices within the context of contemporary mixed-age classrooms. Indeed, much of the discourse surrounding teacher and caregiver stability has been grounded in attachment theory (Ainsworth et al., 1978) and has focused on intra-year changes in caregivers rather than across school years (Phillips, Voran, Kisker, Howes, Whitebook, 1994; Markowitz & Bassok, 2018; Tran & Winsler, 2011). Accordingly, the results of the current investigation extend what is known on this subject matter and indicate that the continuity in caregivers and teachers did *not* have implications for children's early academic achievement (for similar findings see: Horm et al., 2008). Just as importantly, however, our results also revealed that continuity in teachers did *not* minimize the harmful effects of mixed-age classrooms, suggesting that having a larger share of younger classmates in the second year of

Head Start was negatively related to children's academic achievement, and having the same teacher across the two school years did not buffer against this negative association.

Given these findings, there are several implications for both policy and practice. In terms of policy, the results from this investigation suggest that we must pay much closer attention to the ways in which children are placed in classrooms. Given that age cutoffs are likely not feasible in many early childhood programs across the country, one point of intervention might be thinking more carefully about which teachers are assigned to classrooms and investments in teacher professional development. For example, professional development programs that target teachers ability to differentiate their classroom practices is likely to be critical because the vast majority of early childhood educators struggle with this dimension of teaching (Manship et al. 2016). At the same time, however, it is also important to acknowledge that a recent evaluation of a professional development coaching intervention (MyTeachingPartner; Pianta, Mashburn, Downer, Hamre, & Justice, 2008) found that the intervention was ineffective in age diverse settings (Ansari & Pianta, 2018b), which means that professional development opportunities likely need to be tailored to the specific needs of teachers in mixed-age classrooms. An alternative focus might be on teacher education, which a prior study found can mitigate some (but not all) of the negative effects of mixed-age classrooms (Purtell & Ansari, 2018).

Beyond professional development and educational opportunities for teachers, and as briefly discussed above, other aspects of the classroom ecology also require careful attention. For example, in the presence of wide age and skill variability, teachers may use small group instruction to more effectively target their students' needs. Teachers can use these grouping strategies as a means of exposing children to peers of the same age and skill level (i.e., homogeneous groupings), or alternatively, grouping children of varying ages and skill levels (i.e., heterogeneous groupings), both of which have the potential to shape children's academic

learning (Chien et al., 2010; Connor, Morrison, & Slominski, 2006).

Despite these contributions to the early childhood and educational literatures, the results of this investigation need to be interpreted in light of its limitations. First, we were limited in our ability to pinpoint child age to the month, as we relied on teacher reports. Furthermore, because the teachers reported on children's age in the fall in the first year and the spring of the second year, our measure is not parallel across the two years. Second, even though we took several precautions to address concerns regarding omitted variable bias, caution is warranted when interpreting these findings as unmeasured sources of bias may still exist. The best test of the academic impacts of mixed-age education would be randomly assigning children to classrooms of different ages across multiple years, which might not be feasible. Additionally, the small sample size for some of the classroom transitions limited our ability to detect significant differences when effect sizes were smaller, which points to the importance of large-scale data collection efforts that allow for a careful examination of children's educational experiences across multiple school years. However, in our analyses of classroom transitions, we estimated several different thresholds and our findings were largely the same across the various specifications, which tempers this limitation. And even though this study focused on children's academic achievement, there are other dimensions of children's development that require attention, such as their social competence and executive functioning. Unfortunately, however, the social behavior measures in the FACES 2009 data were based on teacher report, which has many limitations, especially in the context of mixed-age classrooms where these reports of individual children are likely made relative to children's classmates.

Moreover, given the high percentage of children from Hispanic and Spanish speaking backgrounds in early childhood programs in the United States, it is also important for future studies to consider how these children experience mixed-age classrooms as compared with their

English-speaking classmates. Finally, the external validity of our findings is limited and is only applicable to Head Start programs, which is why continued research is necessary on the experiences of children in other types of early childhood programs from across the country. It is certainly possible that educational models that place an explicit emphasis on the mixing of ages (e.g., the Montessori Method; Lillard, 2016) result in more optimal outcomes than those documented herein which involved mixed-age classrooms that were not based in pedagogy.

With these limitations and future directions in mind, the present investigation pushes the early childhood field forward and advances our knowledge about age composition in Head Start by providing insight into the ramifications of mixed-age classrooms across multiple school years. Considering that many 3-year-olds who attend Head Start remain in the program for a second year (Puma et al., 2010), our results indicate that we must pay closer attention to children's classroom experiences across school years in order to ensure that they receive enriched educational experiences prior to the transition to kindergarten. Otherwise, children may not reap the maximum benefit possible from Head Start and enter kindergarten at a disadvantage. At the end of the day, despite the fact that mixed-age classrooms have been promoted as a successful educational strategy, the results of this study suggest that caution is warranted when implementing these types of classrooms and that continued empirical inquiry is necessary before firm conclusions can be drawn about the benefits and drawbacks of mixed-age education.

References

Ansari, A. (2017). Multigrade kindergarten classrooms and children's academic achievement, executive function, and socioemotional development. *Infant and Child Development, 26,* e02036. doi: 10.1002/icd.2036

Ansari, A., & Pianta, R. C. (2018a). Teacher-child interaction quality as a function of classroom age diversity and teachers' beliefs and qualifications. *Applied Developmental Science.* Advance online publication. doi: 10.1080/10888691.2018.1439749

Ansari, A., & Pianta, R. C. (2018b). Effects of an early childhood educator coaching intervention on preschoolers: The role of classroom age composition. *Early Childhood Research Quarterly, 44, 101-113.* doi: 10.1016/j.ecresq.2018.03.001

Ansari, A., Purtell, K. M., & Gershoff, E. T. (2016). Classroom age composition and the school readiness of 3- and 4-year-olds in the Head Start program. *Psychological Science, 27,* 53-63. doi: 10.1177/0956797615610882

Administration for Children and Families. (2009). *Head Start Performance Standards*. 45 CFR Ch. XIII (10-1-09 Edition) [With Part 1307, Federal Register Version]. Washington, DC: U.S. Department of Health and Human Services.

Ainsworth, M. D. S., Blehar, M. C., Waters, E., & Wall, S. N. (1978). *Patterns of attachment: A psychological study of the strange situation*. Hillsdale, NJ: Erlbaum.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory.* Englewood Cliffs, NJ: Prentice Hall.

Bell, E. R., Greenfield, D. B., & Bulotsky-Shearer, R. J. (2013). Classroom age composition and rates of change in school readiness for children enrolled in Head Start. *Early Childhood Research Quarterly, 28,* 1-10. doi:10.1016/j.ecresq.2012.06.002

Blasco, P. M., Bailey, D. B., & Burchinal, M. A. (1993). Dimensions of mastery in same-age and mixed-age integrated classrooms. *Early Childhood Research Quarterly*, *8,* 193-206. doi:10.1016/S0885-2006(05)80090-0

Bradbury, B., Corak, M., Waldfogel, J., & Washbrook, E. (2011). *Inequality during the early years: Child outcomes and readiness to learn in Australia, Canada, United Kingdom, and United States* (Institute for the Study of Labor Discussion Paper No. 6120). Retrieved from http://www.econstor.eu/ bitstream/10419/58643/1/690078234.pdf

Bronfenbrenner, U., & Morris, P.A. (2006).The bioecological model of human development. In W. Damon, & R.M. Lerner (Eds.), *Handbook of child psychology (6th ed.). Theoretical models of human development, Vol. 1.* (pp.793– 828). New York: JohnWiley

Burchinal, M. R., Lowe Vandell, D., & Belsky, J. (2014). Is the prediction of adolescent outcomes from early child care moderated by later maternal sensitivity? Results from the NICHD Study of Early Child Care and Youth Development. *Developmental Psychology*, *50*, 542-553. doi: 10.1037/a0033709

Chien, N. C., Howes, C., Burchinal, M., Pianta, R. C., Ritchie, S., Bryant, D. M., ... & Barbarin, O. A. (2010). Children’s classroom engagement and school readiness gains in prekindergarten. *Child Development, 81*, 1534-1549. **doi:** 10.1111/j.1467-8624.2010.01490.x

Connor, C. M., Morrison, F. J., & Slominski, L. (2006). Preschool instruction and children's emergent literacy growth. *Journal of Educational Psychology*, *98*, 665-689. doi:10.1037/0022-0663.98.4.665

DeCoster, J., Iselin, A. M. R., & Gallucci, M. (2009). A conceptual and empirical examination of justifications for dichotomization. *Psychological Methods, 14,* 349-366. doi:10.1037/a0016956

Duncan, G. J., & Magnuson, K. A. (2013). Investing in preschool programs. *Journal of Economic Perspectives*, *27*, 109-132. doi.org/10.1257/jep.27.2.109

Duncan, G. J., Jenkins, J. M., Auger, A., Burchinal, M., Domina, T., & Bitler, M. (2015). Boosting school readiness with preschool curricula. *Irvine Networks on Interventions in Development*.

Duncan, G. J., Engel, M., Claessens, A., & Dowsett, C. J. (2014). Replication and robustness in developmental research. *Developmental Psychology, 50*, 2417-2425. doi: 10.1037/a0037996

Dunn, L. M., & Dunn, L.M. (1997). *Peabody Picture and Vocabulary Test, Third Edition. Examiner's Manual and Norms Booklet*. Circle Pines, MN: American Guidance Service.

Elicker, J., & Fortner-Wood, C. (1995). Adult-child relationships in early childhood programs. *Young Children*, *51*, 69-69.

Guo, Y., Tompkins, V., Justice, L., & Petscher, Y. (2014). Classroom age composition and vocabulary development among at-risk preschoolers. *Early Education and Development*, *25*, 1016-1034. doi: 10.1080/10409289.2014.893759

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

Horm, D. M., File, N., Bryant, D., Burchinal, M., Raikes, H., Forestieri, N., ... & Cobo-Lewis, A. (2018). Associations between continuity of care in infant-toddler classrooms and child outcomes. *Early Childhood Research Quarterly*, *42*, 105-118. doi: 10.1016/j.ecresq.2017.08.002

Jenkins, J. M., Farkas, G., Duncan, G. J., Burchinal, M., & Vandell, D. L. (2016). Head Start at ages 3 and 4 versus Head Start followed by state pre-k: Which is more effective?. *Educational*

Evaluation and Policy Analysis, 38, 88-112. doi: 10.3102/0162373715587965

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

Justice, L. M., Logan, J. A., Purtell, K., Bleses, D., & Hogden, A. (2018). Does mixing age groups in early childhood education settings support children's language development?. *Applied Developmental Science*. Advance online publication. doi: 10.1080/10888691.2017.1386100

Katz, L. G., Evangelou, D., & Hartmann, J. A. (1990). *The case for mixed-age grouping in the early years.* Washington, DC: National Association for the Education of Young Children.

Keys, T. D., Farkas, G., Burchinal, M. R., Duncan, G. J., Vandell, D. L., Li, W., ... & Howes, C. (2013). Preschool center quality and school readiness: Quality effects and variation by demographic and child characteristics. *Child Development*, *84,* 1171-1190. doi: 10.1111/cdev.12048

Lillard, A. S. (2016). *Montessori: The science behind the genius*. New York, NY: Oxford University Press.

Manship, K., Farber, J., Smith, C., & Drummond, K. (2016). *Case studies of schools implementing early elementary strategies: Preschool through third grade alignment and differentiated instruction.* Washington, DC: Office of Planning, Evaluation and Policy Development, US Department of Education.

Markowitz, A.M. & Bassok, D. (2018). *Correlates and consequences of teacher turnover in Head Start.* Paper presented at the 43rd Association for Education Finance and Policy Conference.

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*, 686-702. doi: 10.1111/j.1467-8624.2009.01291.x

Mason, D. A., & Burns, R. B. (1996). "Simply no worse and simply no better" may simply be wrong: A critique of Veenman's conclusion about multigrade classes. *Review of Educational Research, 66,* 307-322. doi:10.3102/00346543066003307

McMullen, M. B., Yun, N., Mihai, A., & Kim, H. (2016). Relationship-based care practice within infant-toddler continuity of care from the perspectives of practitioners and families. *Early Education & Development*, *27,* 190-220. doi: 10.1080/10409289.2016.1102016

Miller, P., Votruba-Drzal, E., McQuiggan, M., & Shaw, A. (2017). Pre-K classroom-economic composition and children's early academic development. *Journal of Educational Psychology*, *109*, 149-165. doi: 10.1037/edu0000137

Moiduddin, E., Aikens, N., Tarullo, L. B., West, J., & Xue, Y. (2012). *Child outcomes and classroom quality in FACES 2009.* Washington, DC: Office of Planning, Research, and Evaluation, U.S. Department of Health and Human Services.

Moller, A. C., Forbes-Jones, E., & Hightower, A. D. (2008). Classroom age composition and developmental change in 70 urban preschool classrooms. *Journal of Educational Psychology*, *100*, 741-753. doi: 10.1037/a0013099

Muthén, L. K., & Muthén, B.f O. (2013). *Mplus User's Guide. Sixth Edition*. Los Angeles, CA: Authors.

National Institute of Child Health and Human Development Early Child Care Research Network & Duncan, G. J. (2003). Modeling the impacts of child care quality on children's preschool cognitive development. *Child Development, 74,* 1454-1475. doi: 10.1111/1467-8624.00617

National Survey of Early Care and Education (2012). *Public dataset for center-based provider survey*. Washington, DC: Administration for Children and Families.

Phillips, D., Lipsey, M. W., Dodge, K.A., Haskins, R., Bassok, D., Burchinal, M. R., Duncan, G. J… Weiland, C. (2017). *Puzzling it out: The current state of scientific knowledge on pre-Kindergarten effects. A consensus statement.* Washington, DC: Brookings Institution.

Phillips, D. A., Voran, M., Kisker, E., Howes, C., & Whitebook, M. (1994). Child care for children in poverty: Opportunity or inequity?. *Child Development*, *65*, 472-492.

Pianta, R. C., La Paro, K. M., & Hamre, B. K. (2008). *Classroom Assessment Scoring System (CLASS)*. Baltimore, MD: Paul H. Brookes.

Pianta, R. C., Mashburn, A. J., Downer, J. T., Hamre, B. K., & Justice, L. (2008). Effects of web-mediated professional development resources on teacher–child interactions in pre-kindergarten classrooms. *Early Childhood Research Quarterly*, *23*, 431-451. doi: 10.1016/j.ecresq.2008.02.001

Pratt, D. (1986). On the merits of multiage classrooms. *Research in Rural Education, 3*, 111–115

Puma, M., Bell, S., Cook, R., Heid, C., Shapiro, G., Broene, P., ... & Spier, E. (2010). *Head Start Impact Study. Final Report.* Washington, DC: Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

Purtell, K. M., & Ansari, A. (2018). Classroom age composition and preschooler's school readiness: The implications of classroom quality and teacher qualifications. *AERA Open, 4,* 1-13. doi:10.1177/2332858418758300

Proehl, R. A., Douglas, S., Elias, D., Johnson, A. H., & Westsmith, W. (2013). A collaborative approach: Assessing the impact of multi-grade classrooms. *Catholic Education: A Journal of Inquiry and Practice, 16,* 417-440.

Radloff, L. S. (1977). The CESD scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*, *1*, 385–401. doi:10. 1177/014662167700100306

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. doi.org/10.1016/j.ecresq.2017.05.003

Schafer, J. L., & Graham, J. W. (2002). Missing data: our view of the state of the art. *Psychological Methods*, *7*, 147-177. doi: 10.1037/1082-989X.7.2.147

Snow, K., Thalji, L., Derecho, A., Wheeless, S., Lennon, J., Kinsey, S., Rogers, J. et al. (2007). *ECLS-B: Data file user's manual.* Washington, DC: National Center for Education Statistics.

Sosinsky, L., Ruprecht, K., Horm, D., Kriener-Althen, K., Vogel, C., & Halle, T. (2016). *Including relationship-based care practices in infant-toddler care: Implications for practice and policy*. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

StataCorp. (2009). *Stata user's guide, release 11.* Stata Press: College Station, TX.

Thomas, J. L. (2012). Combination classes and educational achievement. *Economics of Education Review*, *31*, 1058-1066. doi:10.1016/j.econedurev.2012.07.013

Tran, H., & Winsler, A. (2011). Teacher and center stability and school readiness among low-income, ethnically diverse children in subsidized, center-based child care. *Children and Youth Services Review*, *33*, 2241-2252. doi: 10.1016/j.childyouth.2011.07.008

Urberg, K. A., & Kaplan, M. G. (1986). Effects of classroom age composition on the play and social behaviors of preschool children. *Journal of Applied Developmental Psychology, 7,* 403–415. doi:10.1016/0193-3973(86)90009-2

Veenman, S. (1995). Cognitive and noncognitive effects of multigrade and multi-age classes: A best-evidence synthesis. *Review of Educational Research*, *65*, 319-381. doi: 10.3102/00346543065004319

Vygotsky, L. S. (1978). Interaction between learning and development. In M. Cole, V. John-Steiner, S. Scribner, & E. Souberman (Eds.), *Readings on the development of children* (pp. 34–41). Cambridge, MA: Harvard University Press.

Way, J. W. (1981). Achievement and self-concept in multiage classrooms. *Educational Research Quarterly, 6,* 69-75

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

Yoshikawa, H., Weiland, C., Brooks-Gunn, J., Burchinal, M., Espinosa, L. M., Gormley, W. T., . . . Zaslow, M. J. (2013). *Investing in our future: The evidence base on preschool education.* New York, NY: Foundation for Child Development.

Table 1.
Weighted descriptives child and household variables.

Variables	Mean (SD) or proportion
Child gender (proportion female)	0.49
Child race	
White	0.20
Black	0.34
Latino	0.37
Asian/other	0.08
Child age at entry into Head Start (fall 2009)	41.13 (3.68)
Child age at spring 2011	60.16 (3.71)
Months between fall 2009 and spring 2010 assessments	5.69 (1.87)
Months between spring 2010 and spring 2011 assessments	11.53 (0.72)
Language of assessment	
English (fall 2009)	0.74
English (spring 2010)	0.89
English (spring 2011)	0.98
Mothers' marital status	
Married	0.32
Not married	0.17
Not two parent household.	0.51
Mothers' education	
Less than high school	0.35
High school diploma	0.36
Some college	0.22
Bachelor's degree	0.06
Mothers' age (mean years)	28.93 (5.86)
Household size (mean number of people)	4.68 (1.65)
Mothers' employment	
Full time	0.26
Part time	0.21
Unemployed	0.54
Mothers' depressive symptoms	4.96 (5.96)
Ratio of income to poverty	2.52 (1.35)
Household language (English)	0.75
Child outcomes	
Language (fall 2009)	81.33 (19.48)
Language (spring 2010)	85.05 (16.68)
Language (spring 2011)	91.77 (14.87)
Letter word identification (fall 2009)	298.67 (22.48)
Letter word identification (spring 2010)	315.65 (25.88)
Letter word identification (spring 2011)	337.91 (25.99)
Spelling (fall 2009)	331.89 (25.75)
Spelling (spring 2010)	350.62 (27.37)
Spelling (spring 2011)	381.46 (28.18)
Math (fall 2009)	10.89 (4.89)
Math (spring 2010)	15.19 (6.80)
Math (spring 2011)	24.75 (7.74)

Notes. Proportions might not sum to 1.00 due to rounding.

Table 2.
Weighted descriptives for focal classroom variables

	Mean (SD) or proportion	
Variables	Year 1	Year 2
Classroom age composition		
Proportion 3 year olds	0.58	0.14
Proportion 4 year olds	0.42	0.46
Proportion 5 year olds	---	0.39
Classroom transitions: Thirty percent threshold		
Younger different-age to older different-age classroom	---	0.52
Younger different-age to older same-age classroom	---	0.05
Younger same-age to older different-age classroom	---	0.33
Younger same-age to older same-age classroom	---	0.10
Classroom transitions: Fifty percent threshold		
Younger different-age to older different-age classroom	---	0.32
Younger different-age to older same-age classroom	---	0.14
Younger same-age to older different-age classroom	---	0.32
Younger same-age to older same-age classroom	---	0.22
Classroom transitions: Seventy percent threshold		
Younger different-age to older different-age classroom	---	0.09
Younger different-age to older same-age classroom	---	0.16
Younger same-age to older different-age classroom	---	0.27
Younger same-age to older same-age classroom	---	0.49
Same teacher from year 1 to year 2	---	0.36
Child/teacher ratio	8.42 (2.13)	8.66 (1.82)
Child/adult ratio	7.30 (2.16)	7.69 (1.86)
Class size (mean number of children)	16.83 (2.24)	17.82 (2.02)
Teachers' depressive symptoms	4.78 (5.07)	4.16 (5.28)
Hours of school per week	25.18 (11.72)	27.11 (11.55)
Multilingual instruction	0.33	0.29
Teachers' years of teaching experience	13.09 (8.76)	13.02 (9.05)
Teachers' education		
High school	0.05	0.06
Some college	0.12	0.12
Associates	0.35	0.31
Bachelors	0.34	0.37
Some graduate school	0.13	0.13
Teachers' degree in early childhood education	0.93	0.92
Teachers' race		
White	0.41	0.40
Black	0.36	0.37
Latino	0.20	0.17
Asian/other	0.02	0.05
Teachers' hourly salary ($)	14.07 (6.20)	14.02 (6.35)
Teachers' benefits (scale from 0-9)	6.79 (2.13)	6.95 (1.97)
Teacher teaches multiple classes	0.16	0.15
Teacher-child interaction quality	4.11 (0.50)	4.13 (0.54)

Notes. Proportions might not sum to 1.00 due to rounding.

Table 3.
Results of models predicting children's early academic achievement.

	Reading				Math			
Variables	Model 1	Model 2	Model 3	Model 4	Model 1	Model 2	Model 3	Model 4
Proportion of younger peers during year 2	-0.07 * (0.03)		0.07 * (0.03)	-0.06 † (0.03)	-0.05 * (0.02)		-0.05 * (0.02)	-0.03 (0.03)
Proportion of 3 year olds during year 2		-0.08 ** (0.03)				-0.04 (0.03)		
Proportion of 4 year olds during year 2		-0.03 (0.03)				-0.04 (0.03)		
Proportion of same age peers during year 1	0.00 (0.03)	-0.01 (0.03)	0.00 (0.03)	0.00 (0.03)	0.01 (0.03)	0.01 (0.03)	0.01 (0.03)	0.01 (0.03)
Stability in teacher from year 1 to year 2	0.04 (0.06)	0.05 (0.06)	0.04 (0.06)	0.05 (0.06)	-0.02 (0.05)	-0.02 (0.06)	-0.02 (0.06)	-0.00 (0.06)
Classroom age composition in year 1 X year 2			0.04 (0.03)				0.04 (0.03)	
Proportion of younger peers during year 2 X teacher stability				-0.03 (0.06)				-0.07 (0.05)

Notes. All continuous variables were standardized (mean of 0 and standard deviation of 1), and therefore the unstandardized regression coefficients in this table correspond to effect sizes (i.e., standard-deviation units). $R^2 = 0.61$ and 0.64 for reading and math, respectively. Models adjusted for the clustering of children in classrooms and all covariates listed in Tables 1 and 2. *** $p < .001$. ** $p < .01$. * $p < .05$. † $p < .10$.

Table 4.
Results of transition models predicting children's early academic achievement.

	Thirty percent threshold					
	Younger different-age peers Older different-age peers		Younger different-age peers Older same-age peers		Younger same-age peers Older different-age peers	
	Math	Reading	Math	Reading	Math	Reading
Younger different-age peers	---	---				
Older different-age peers						
Younger different-age peers	0.31 **	0.38 **	---	---		
Older same-age peers	(0.11)	(0.14)				
Younger same-age peers	0.11 †	0.12 †	-0.20 †	-0.26 †	---	---
Older different-age peers	(0.06	(0.07)	(0.12)	(0.14)		
Younger same-age peers	0.14	0.11	-0.18	-0.26	0.03	-0.00
Older same-age peers	(0.11)	(0.13)	(0.15)	(0.17)	(0.10)	(0.12)
	Fifty percent threshold					
	Younger different-age peers Older different-age peers		Younger different-age peers Older same-age peers		Younger same-age peers Older different-age peers	
	Math	Reading	Math	Reading	Math	Reading
Younger different-age peers	---	---				
Older different-age peers						
Younger different-age peers	0.17 *	0.15 †	---	---		
Older same-age peers	(0.07)	(0.08)				
Younger same-age peers	0.06	0.09	-0.11	-0.07	---	---
Older different-age peers	(0.07)	(0.07)	(0.08)	(0.08)		
Younger same-age peers	0.16 *	0.11	-0.01	-0.04	0.10	0.03
Older same-age peers	(0.08)	(0.09)	(0.09)	(0.10)	(0.07)	(0.08)
	Seventy percent threshold					
	Younger different-age peers Older different-age peers		Younger different-age peers Older same-age peers		Younger same-age peers Older different-age peers	
	Math	Reading	Math	Reading	Math	Reading
Younger different-age peers	---	---				
Older different-age peers						
Younger different-age peers	0.23 *	0.24 *	---	---		
Older same-age peers	(0.10)	(0.11)				
Younger same-age peers	0.10	0.06	-0.13 †	-0.18 *	---	---
Older different-age peers	(0.10)	(0.10)	0.08)	(0.09)		
Younger same-age peers	0.11	0.11	-0.12	-0.13	0.02	0.05
Older same-age peers	(0.09)	(0.09)	(0.07)	(0.08)	(0.06)	(0.07)

Notes. Columns are the referent group. Focal predictors are the rows. All continuous variables were standardized (mean of 0 and standard deviation of 1), and therefore the regression coefficients in this table correspond to effect sizes (i.e., standard-deviation units). Models adjusted for the clustering of children in classrooms and all covariates listed in Tables 1 and 2. *** $p < .001$. ** $p < .01$. * $p < .05$. † $p < .10$.