Submitted *Developmental Psychology*, August 9, 2017
Revision #1 December 16, 2017
Accepted December 31, 2017

VARIATION IN THE LONG-TERM BENEFITS OF CHILD CARE: THE ROLE OF CLASSROOM QUALITY IN ELEMENTARY SCHOOL

Arya Ansari
University of Virginia

Robert Pianta
University of Virginia

Running Head: Variation in Child Care Benefits
Word Count: 127 (abstract), 8,923 (focal text),
5 Tables, 0 Figures, 9 Appendices (4 tables and 5 figures)

Ansari, A., & Pianta, R. C. (in press). Variation in the long-term effects of child care: The role of classroom quality during middle childhood. *Developmental Psychology*. Advance online publication. doi: 10.1037/dev0000513

* 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 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

Data from the NICHD Study of Early Child Care and Youth Development (*n* = 1,307) were used to estimate the additive and multiplicative benefits of high quality child care between birth to 54 months of age and high quality elementary school education between first and fifth grade. Results indicated that the math and language and literacy benefits of high quality child care accrued from the end of preschool through age 15 only when coupled with higher quality classroom environments during the elementary school years. In contrast, the benefits of high quality child care were no longer present when children later attended lower quality classrooms in elementary school. Taken together, these results point to the importance of continued investments in children's education throughout the first decade of life.

Keywords: Child care quality; classroom quality; academic achievement; NICHD SECCYD

Variation in the Long-term Benefits of Child Care:
The Role of Classroom Quality in Elementary School

There is strong research and policy interest in the early childhood years, and in early care and education programs in particular, as a potential remedy for long-term inequality (Phillips et al., 2017; Yoshikawa et al., 2013). Indeed, recent national estimates from the United States suggest that approximately half of infants and toddlers under 4 years of age experience regular non-parental care and up to 75% of 4-year-olds experience child care during the year before kindergarten (Burchinal, Magnuson, Powell, & Hong, 2015). Although it is often argued that early investments are one of the most cost-effective means of improving children's long-term educational prospects (Heckman, 2006), the efficacy of child care and other early childhood programs are increasingly in question. This is because the long-term impacts of contemporary early childhood programs have been ambiguous, with some scholars documenting continued, albeit modest, benefits throughout middle childhood and adolescence (Ansari et al., 2016; Ansari, in press; Muschkin et al., 2015; Vandell, Burchinal, & Pierce 2016) and others finding no discernable effects shortly after kindergarten (Hill et al., 2015; Lipsey et al., 2015; Magnuson et al., 2007; Puma et al., 2012).

In view of these conflicting claims, a pressing question in the developmental sciences has centered on these patterns of convergence and persistence in the benefits of early childhood programs and in identifying the conditions under which the effects of early experiences are accentuated, attenuated, or maintained (Duncan & Magnuson, 2013; Yoshikawa et al., 2013). To this end, although some scholars find that children's experiences in middle childhood and adolescence override any experiences in early childhood and thus should be considered more salient developmentally (Bradley, Caldwell, & Rock, 1988), others contend that children's post-

preschool experiences play a more nuanced role (Bailey, Duncan, Odgers, & Yu, 2016). More specifically, one of the arguments put forth by Bailey and colleagues (2016) is that the legacy impact of early experiences vary as a function of experience in subsequent exogenous environments. This "sustaining environments" argument has received growing empirical attention, with recent calls for initiatives that provide children with continued high quality educational opportunities across the first decade of life (Reynolds & Temple, 2014). That is, despite the fact that investments in early childhood programs can (and do) prepare children for kindergarten (e.g., Ansari, in press; Fuller, Bein, Bridges, Kim, & Rabe-Hesketh, 2017; Gormley, Phillips, & Gayer, 2008; Magnuson et al., 2007; Puma et al., 2012; Weiland & Yoshikawa, 2013), children's long-term school success is likely a function of concatenated investments that contribute to sustaining or fading impacts of early experiences, including those in child care programs.

Although intuitive, the empirical evidence in support of this position is fairly limited, in large part because studies of early care and education programs rarely gather data on children's experiences in subsequent educational settings. Yet, even among the handful of studies that have considered different aspects of the elementary school environment as potential moderators of the long-term effects of child care (e.g., Bassok, Gibbs, & Latham, 2015; Claessens, Engel, & Curran, 2014; Currie & Thomas, 2000; Lee & Loeb, 1995; Zhai, Raver, & Jones, 2012), the evidence in support of the sustaining environments argument is equally mixed. For example, some educational scholars found that school environments of sufficient quality do sustain the benefits of early investments in children's education (Currie & Thomas, 2000; Lee & Loeb, 1995; Zhai et al., 2012). In contrast, others found that the long-term benefits of early care and education programs are *not* conditional on children's later schooling and, ultimately, the benefits

conferred by these programs diminish to a similar degree for *all* children as they progress throughout the elementary school years (Bassok et al., 2015; Claessens et al., 2014).

To date, however, much of the educational literature that has tested the evidence for the sustaining environments argument (Bailey et al., 2016), at least with respect to the school environment, has frequently examined structural measures of classroom quality, such as elementary program type and class size as opposed to process quality, such as teachers' day-to-day interactions with their students. Considering that contemporary studies of education and development emphasize the role of children's proximal interactions with their teachers as a powerful force in shaping children's short- and long-term school success (Pianta & Hamre, 2009), it is likely that such interactions play a stronger role in the persistence (or convergence) of child care effects as compared with other aspects of schooling, such as class size. And it should be recognized that even with growing investments in child care and in early education programs across the country, there is not a coherent approach to providing continuous (i.e., year after year) high quality educational opportunities for young children in the United States (Pianta, Belsky, Houts, & Morrison, 2007); therefore, it should not be at all surprising that even the most robust early childhood effects diminish over time (e.g., Hill et al., 2015; Lipsey et al., 2015; Puma et al., 2012).

It is also important to acknowledge that the sustaining environments hypothesis, which points to the conditional nature of early childhood effects—although the focus of this investigation—is but one of the many theories behind the convergence (or maintenance) of the benefits that emerge from early intervention and child care programs. For example, economic theories of skill building suggest that early investments shape children's long-term development by providing children with the foundational skills necessary to succeed in school (Cunha et al.,

2006). As one example, children's early counting skills have been documented as the basis for math (and reading) achievement in subsequent years such that children with higher math abilities during the early years are more likely to do better on more advanced assessments of mathematics knowledge (Duncan et al., 2007). Thus, programs that boost children's math achievement, in theory, may have sustained long-term benefits for children academically regardless of subsequent experiences. At the same time, however, Bailey and colleagues (2016) argue that the focal skills children learn matters: these skills have to be *malleable* to intervention, *fundamental* for later success, and would *not otherwise develop over time* in the counterfactual. A third possibility stems from the "foot in the door" hypothesis (Bailey et al., 2016), which argues that successful early childhood programs are those that equip children with the right skills to succeed and avoid imminent risks (e.g., being held back in school) or to seize other opportunities (e.g., entry into honors classes), which in turn have long-term cascading effects for children.

Notwithstanding these other arguments for variation in the longer-term benefits of early care and education programs, in the current investigation we focus on the sustaining environments hypothesis. More specifically, we extend the existing literature by examining the degree to which the long-term benefits of child care quality are conditioned on children's later experiences in school, with a focus on children's academic learning, as they relate to the quality of their proximal experiences in the classroom during the early and middle childhood years. As part of this effort, we consider both children's cumulative educational experiences during the elementary school years and the specific years during which they experience higher quality classroom environments. We consider both the cumulative and time-specific nature of children's experiences because, developmentally speaking, the benefits of child care and educational quality for children may increase from exposure over time, which would be captured with our

cumulative model, whereas the timing specific models could point to periods during which quality makes a larger difference. And although a timing specific model might indicate when quality matters most, if the benefits of child care or classrooms at any given point in time are weak (e.g., Keys et al., 2013) or inconsistent over time (e.g., Pianta et al., 2007), then it might be children's cumulative experiences in the classroom that result in significant effects rather than those from any one given period.

When taken together, although previous studies with the National Institute of Child Health and Human Development (NICHD) Study of Early Child Care and Youth Development (SECCYD) indicate that child care quality predicts higher academic test scores throughout the first 15 years of life (Belsky et al., 2007; Burchinal et al., 2014; Vandell et al., 2010), this is the first to consider long-term variation in the persistence of these associations as a function of children's subsequent experiences in the classroom. Drawing from the sustaining environments model (Bailey et al., 2016), we hypothesize that the benefits of high quality child care will be more pronounced over time when children subsequently experience higher quality education as opposed to lower quality education, where the initial academic advantages of child care will be more likely to diminish as children progress through the K-12 educational pipeline.

Method

Data for the current investigation were drawn from the NICHD SECCYD, a longitudinal investigation of children from birth through the end of high school. Potential participants were pooled from mothers who gave birth during selected sampling periods in 1991 across 10 different sites in the United States (Little Rock, AR; Irvine, CA; Lawrence, KS; Boston, MA; Philadelphia, PA; Pittsburgh, PA; Charlottesville, VA; Morganton, NC; Seattle, WA; and Madison, WI). A total of 1,364 mothers with healthy newborns were enrolled into the study and

participated in Phase I (ages 0-3; 1991-1994). Of these 1,364 study participates, 1,226 participated in Phase II (through first grade; 1995-1999), 1,061 participated in Phase III (through sixth grade; 2000-2004), and 1,009 participated in Phase IV (through ninth grade; 2005-2007). As part of the longitudinal investigation, data were regularly collected from multiple informants including interviews with parents and teachers, direct assessments of children, and observations of the home and school. The analytic sample for the current investigation includes 1,307 children and families of the original recruitment sample who were *not* exclusively cared for by their mothers. Thus, our study sample includes children who experienced nonmaternal care at some point between birth and 54 months of age. Participating children were largely White (81%), had mothers who averaged a little over 14 years of education, and lived in households between birth to 54 months of age that had an average family income that was approximately 3.64 times greater than the poverty threshold. Because this investigation is based on secondary data that are available through the Interuniversity Consortium for Political and Social Research (ICPSR), our research ethics committee at the University of Virginia declared this study exempt.

Measures

For descriptive statistics for all focal variables and covariates, see Tables 1 and 2. For a correlation matrix of the focal variables of interest, see Supplemental Table 1.

Child care quality. Child care quality was based on observational assessments in children's primary child care arrangement at 6, 15, 24, 36, and 54 months of age. Observational assessments of caregiver-child interactions were obtained using the Observational Record of the Caregiving Environment (ORCE) for children who were in 10 or more hours per week of nonmaternal care. These assessments occurred across two half-day visits within a two-week period when children were 6 to 36 months of age and one half-day visit at 54 months of age, all

of which demonstrated strong internal consistency (α = .90, .86, .81, .80, .90 at 6, 15, 24, 36, and 54 months, respectively). Before conducting observations, all coders received extensive training and, to be certified, coders had to achieve exact agreement with the master codes of the qualitative ratings at 60% or better. Inter-observer agreement exceeded 0.80 across all ages.

Once certified, the trained observers completed four 44-min cycles of the ORCE at 6, 15, 24, and 36 months and two 44-min cycles at 54 months. These observations captured caregivers' sensitivity to children's non-distress signals, stimulation of children's development, positive regard toward children, detachment, flatness of affect, and intrusiveness. These scores were based on a 4-point Likert scale (1 = *not at all characteristic*, 4 = *high characteristic*). Prior studies with this instrument have suggested that an overall composite is most appropriate and, thus, these ratings were averaged to create a composite of child care quality spanning from birth to 54 months of age (see also: Belsky et al., 2007; Burchinal et al., 2014; NICHD Early Child Care Research Network [ECCRN] & Duncan, 2003; NICHD ECCRN, 1997, 2002; Vandell et al., 2010, 2016). At least one observational assessment was available for 86% of children between 6 and 54 months (45%, 50%, 51%, 54%, and 65% at 6, 15, 24, 36, and 54 months).

Classroom quality during elementary school. Classroom quality during the elementary school years was assessed with the Classroom Observation System (COS) during first (COS-1), third (COS-3), and fifth grade (COS-5). These classrooms observations included global ratings of teacher-child interactions and occurred during the start of the school day and took roughly three hours for the COS-1 and six hours for the COS-3 and COS-5. These global ratings captured several dimensions of process quality: classroom overcontrol, chaos, positive emotional climate, negative emotional climate, teacher detachment, teacher sensitivity, productive use of instructional time, and rich instructional methods (see also: Pianta, Belsky, Vandergrift, Houts,

& Morrison, 2008), with ratings ranging from 1 (*uncharacteristic*) to 7 (*extremely characteristic*). For each grade level, we created a composite of classroom quality by averaging the COS ratings for that grade (α = .88, .79, .82 at first, third and fifth grade, respectively). In total, 83% of the study children had at least one classroom observation in elementary school (73%, 73%, 72% at first, third, and fifth grade).

Before rating classrooms, however, all coders underwent extensive reliability training and testing. To be deemed reliable, 80% of a coder's ratings needed to be within 1 scale point of the master ratings. All coders met or exceeded this level of reliability. Additionally, Pianta and colleagues (2008) examined the relation between codes for classrooms observed more than once in the NICHD SECCYD to estimate the stability of the observations across days and across different children in first (n = 63 classrooms), third (n = 52 classrooms), and fifth (n = 54 classrooms) grade. Results from this assessment—a form of test-retest reliability—revealed an average cross-day correlation of .71, .88, and .91 in first, third, and fifth grade, respectively, Thus, these ratings are reflective of aspects of the classroom ecology that remain relatively stable across time and children.

It is important to note that, unlike the ORCE, the COS can be used either as an overall composite or as a two-factor structure to capture instructional and emotional dimensions of the classroom. Although our focal analyses use the overall composite of classroom quality during the elementary school years, we estimate supplemental models looking at the main and moderated effects of instructional and emotional support as separate indicators. The results from these models indicate that the associations between these two dimensions of the COS and children's school success are statistically indistinguishable from one another (see Model 1 of Supplemental Table 2) and support the general conclusions discussed below.

Academic achievement. Children's academic achievement was directly assessed at 54 months of age, first grade, third grade, fifth grade, and ninth grade with the Woodcock-Johnson Educational Battery–Revised (Woodcock & Johnson, 1989), which included assessments of children's language and literacy and math skills. Among the full sample, roughly 80%, 77%, 76%, 75%, and 67% of children had valid test scores at 54 months of age, first, third, fifth, and ninth grade, respectively. For the purposes of the current study we created a composite for language/literacy (letter word identification, word attack, and passage comprehension; $\alpha = .60$, .78, .89, .84, .83 at 6, 15, 24, 36, and 54 months, respectively) and math (applied problems). Given concerns surrounding missing data and sample attrition, we also estimated models among the sample of children who participated through age 15 and all findings were qualitatively and quantitatively similar to those reported below (see Model 2 of Supplemental Table 2).

Maternal, child, and family confounds. One of the major concerns with early intervention research and studies on education quality is that the focal predictor—in this case child care quality and classroom quality in elementary school—is endogenous. That endogeneity can undermine causal inferences to be made about associations between education quality and children's academic achievement, as factors that select children into higher quality environments might also influence their school success (Duncan & Magnuson, 2013; NICHD ECCRN & Duncan, 2003). To address these issues we explicitly estimate factors selecting children into different quality environments and identify measureable confounds that need to be controlled when examining the additive and multiplicative benefits of education quality for children's school performance. Each of these factors, in turn, is included in regression models predicting children's math and language and literacy achievement. These variables included are largely based on prior child care studies done with the NICHD SECCYD (e.g., Belsky et al., 2007;

Dearing et al., 2009; NICHD ECCRN & Duncan, 2003; Vandell et al., 2010, 2016) as well as theoretically informed studies done on child care and preschool selection (e.g., Ansari, 2017; Crosnoe, Purtell, Davis-Kean, Ansari, & Benner, 2016; Coley et al., 2014).

Before discussing these factors, however, it is important to note that similar to prior child care reports with these data (e.g., Belsky et al., 2007; Burchinal et al., 2014; Vandell et al., 2010, 2016), when available, we also adjusted for time varying covariates for birth to 54 months (6, 15, 24, 36, 54 months of age), the elementary school years (first, third, and fifth grade), and adolescence (ninth grade) as a means of providing a more conservative assessment of the effects of early experiences on long-term development. If, however, early child care affected the covariates during the later years (i.e., the elementary school years and adolescence), then their inclusion in our models might result in a biased estimate of the effects of child care and education quality. For these reasons, we also estimated a series of models without covariates from the elementary school years and adolescence, which resulted in the same pattern of results as those discussed below (see Model 3 of Supplemental Table 2).

To begin, at the child-level, all of our models adjusted for children's: race (White, Black, other), birthweight (in pounds), birth order, mother report of children's temperament (Infant Temperament Questionnaire; Medoff-Cooper, Carey, & McDevitt, 1993) and health (1= *poor*, 4= *excellent*) before entering child care (i.e., at 6 months of age), and the proportion of time children attended center care between birth to 54 months. All models also adjusted for children's attachment security at 36 months, which was assessed with a modified version of the Strange Situation (Cassidy et al., 1992) and classified children as either securely or insecurely attached (see also: Dallaire & Weinraub, 2007; McCartney et al., 2004; NICHD ECCRN, 2001b).

At the parent level, all models adjusted for a wide variety of indicators, both time varying

and time invariant. Time invariant factors included: mothers' age after birth of child, mothers' receipt of maternity leave (0 = *no,* 1 = *yes*), mothers' report of their own health when children were 6 months of age (1= *poor*, 5= *excellent*), mothers' years of education at birth of child, mothers' vocabulary skills at 36 months of age (measured with the Peabody Picture Vocabulary Test-Revised; Dunn & Dunn, 1981), mothers' psychological adjustment when children were 6 months of age (NEO Personality Inventory; 55 items, $\alpha = 0.81$; Costa & McCrae, 1985, 1989), and mothers' childrearing attitudes after birth of child, including their progressive (8 variables, $\alpha = 0.60$; e.g., children should be allowed to disagree with their parents) and traditional (22 items, $\alpha = 0.90$; e.g., preparing for the future is more important for children than enjoying today) beliefs (Parental Modernity Scale of Childrearing and Educational Beliefs; Schaefer & Edgerton, 1985). After the birth of their child, mothers also reported how they felt when they found out they were pregnant (1= very *unhappy*, 5 = *happy*), how they felt about their baby after birth (0 = *mixed feelings*, 1= *happy*), and their locus of control (20 items, $\alpha = 0.61$; e.g., being a good parent often depends on being lucky enough to have a good baby). All models also adjusted for site fixed effects (i.e., Little Rock, AR; Irvine, CA; Lawrence, KS; Boston, MA; Philadelphia, PA; Pittsburgh, PA; Charlottesville, VA; Morganton, NC; Seattle, WA; and Madison, WI).

As noted above, we also included a series of time varying covariates across birth to 54 months, the elementary school years, and adolescence, namely: the proportion of time mothers were employed, the proportion of time mothers were enrolled in school, the average household income-to-needs ratio, the proportion of time children lived in a two-parent household, household size, mothers' depressive symptoms (measured with the Center for Epidemiological Studies Depression Scale; Radloff, 1977), mothers' relationship quality with their partner (26 items, $\alpha = .90$; e.g., my spouse/partner listens to me when I need someone to talk to), parenting

quality as measured by the Home Observation for Measurement of the Environment Scale (e.g., availability of learning materials, language stimulation, responsivity; Caldwell & Bradley, 1984), along with maternal sensitivity from videotaped interactions (see also: Burchinal et al., 2014; NICHD ECCRN, 2002; Vandell et al., 2010, 2016; Watamura et al., 2011). Finally, our models also adjusted for principal reports of broader school quality (19 items, αs = .88-.93; e.g., problems with child absences, robbery and theft, drug use) and a composite of neighborhood disadvantage, which was based on census data and captured the percent of: unemployed adults, adults with less than a high school degree, households living in poverty, and households receiving government assistance.

Analytic Strategy. Our focal analyses were estimated within a regression framework in the *Mplus* program (version 7; Muthén & Muthén, 1998-2013) and missing data, which ranged from 0-42% (mean = 15%) were accounted for with full information maximum likelihood estimation (FIML; Schafer and Graham, 2002). Because our analyses focused on children's child care experiences across birth to 54 months of age, there is considerable variation in dosage and the proportion of time children spent in non-maternal care during this period. On the one hand, controlling for variation in dosage would treat the quality of, for example, one year in non-maternal care the same as, for example, nearly 3 years of non-maternal care with the expectation that these would have comparable effects. On the other, not adjusting for these differences would give undue weight to children who experienced *less* time in non-maternal care. Thus, to address this issue, we weighted our models by the proportion of time children spent in non-maternal care. In doing so, children who spent more time in non-maternal care received greater weight than children who spent less time in non-maternal care between birth to 54 months of age.

With the above in mind, our focal analyses were completed in two steps. In our first set

of analyses, we estimated an omnibus model in which we examined the role of child care quality between birth to 54 months of age (i.e., a combination of child care quality at 6, 15, 24, 36, and 54 months of age) and classroom quality during the elementary school years (i.e., combination of classroom quality during first, third, and fifth grade). It is important to note that the omnibus indicator of education quality during the elementary school years was aligned to the grade at which the outcome assessment was available during that larger time frame. For example, when looking at children's third grade academic test scores, the omnibus indicator of classroom quality only included children's experiences in first and third grade, whereas when looking at fifth and ninth grade outcomes, the composite of classroom quality included children's experiences in first, third, and fifth grade. That is, prospective indicators of classroom quality during the elementary school years were *not* included in our models of past academic performance.

Next, to test for moderation, we interacted the focal independent variables (i.e., child care quality X classroom quality during elementary school). If there was evidence for moderation, we plotted the interactions by calculating the predicted outcome scores for a one standard deviation change in child care quality in low (i.e., one standard deviation below the mean) and high (i.e., one standard deviation above the mean) quality classrooms during the elementary school years (Aiken & West, 1991). In our second series of analyses, we estimated timing specific models examining the contribution of classroom quality during first, third, and fifth grade, separately. A similar series of interactions were also estimated for our timing specific analyses. Finally, it is important to note that although all models included the covariates listed in Table 1, we did not include prospective covariates (e.g., covariates from adolescence) when estimating models for earlier time points (e.g., the elementary school years).

Results

We begin with descriptive statistics for our focal variables. We then summarize the main effects of child care quality between birth and 54 months and classroom quality during the elementary school years before turning to a discussion of the conditional nature of the long-term benefits of child care quality as a function of children's subsequent classroom experiences. We close with a presentation of robustness checks and a series of models that compare the conditional nature of child care quality as a function of: (a) children's cumulative experiences in elementary school; and (b) the specific years in which they experience higher quality classrooms.

Descriptives and bivariate correlates of education quality. Between birth and 54 months of age, 70% of the time children were sampled they were *not* exclusively cared for by their mothers (*SD* = 26%, range =6-100%). Not surprisingly, when children were one month of age 93% were cared for exclusively by their mothers, but this steadily dropped to 34%, 29%, 27%, 20%, and 10% at 6, 15, 24, 36, and 54 months of age and, given the sample restrictions, all children experienced non-maternal care at some point between birth and 54 months of age. During this same time frame, the study children spent approximately 20% (*SD* =26%, range 0-94%) of their time in center-based care for 10 or more hours per week. Put another way, for roughly a third of the time children were *not* cared for by their mothers between 6 and 54 months of age, they were enrolled in a center-based program.

Moreover, between 6 and 54 months of age, children experienced moderate quality care (*M* = 2.92, *SD* = 0.42; for a histogram see Supplemental Figure 1), with roughly 19% attending child care programs that were considered to be of high quality (i.e., above a 3.30; thresholds of child care quality are based on the work of Vandell et al., 2010) and 33% experiencing care that was of low quality (i.e., scores below 2.75). Children subsequently attended elementary school classrooms that were of moderate to moderately high quality in first (*M* = 5.10, *SD* = 0.86), third

(*M* = 5.25, *SD* =0.56), and fifth (*M* = 5.23, *SD* = 0.54) grade (see Supplemental Figures 2-5 for histograms), with 27% attending classrooms that were of high quality (score above 5.50). As can be seen by the bivariate correlations (see Supplemental Table 1), however, children who experienced higher quality interactions during early childhood were not much more likely to experience a higher quality education in elementary school (r = 0.15, p < .001) and children who experienced higher quality education in any grade level during elementary school were not guaranteed to experience higher quality education during later grades (rs = 0.14-0.22). In fact, and as can be seen in Table 3, roughly 23% of children who experienced low quality child care (7% of the full sample) went on to experience high quality elementary school education, whereas another 22% of children who experienced high quality child care (4% of the full sample) subsequently experienced low quality elementary school education. Thus, overall, only a small fraction of children experienced consistent high quality education from birth through fifth grade.

Despite the fact that children infrequently experienced consistent high quality education across the first decade of life, both bivariate and multivariate models revealed that some children were more likely to experience higher quality education than others. Below we only discuss the multivariate differences, but bivariate differences between the early childhood demographics and child care and elementary school quality are also presented in Table 4. To begin, we found that children who lived in larger homes and whose mothers were employed and received maternity leave and children who spent greater time in center care between 6 and 54 months of age experienced *lower* quality child care. In contrast, children who lived in more affluent homes and children who experienced higher quality parenting and had healthier mothers experienced higher quality child care. In terms of classroom quality in elementary school, results suggested that boys, black children, and children born to younger mothers experienced lower quality education,

whereas children who lived in homes that were more affluent, those who experienced greater maternal sensitivity, and those whose mothers attended school for more epochs attended classrooms of higher quality. Effect sizes for these associations ranged from 8-30% of a *SD*. Outside of these factors, however, the remaining child and family characteristics were not significantly associated with the quality of children's experiences in child care or elementary school and the average bivariate (*r*s = ~0.13-0.14) and partial (*r*s = ~0.05-0.06) correlations between all of the covariates and child care and elementary school quality indicators was small.

Main effects of child care and classroom quality. Having examined both the descriptive patterns of child care and elementary school quality and considered which children were more (or less) likely to experience high quality education, we next estimated a series of regression models that examined the main effects of child care quality and classroom quality in elementary school for children's math and language and literacy skills from 54 months through age 15. For these main effects models, we include all quality variables as predictors along with the covariates. These analyses were also weighted by the proportion of time children spent in non-maternal care.

Similar to prior studies done with the NICHD SECCYD on children's academic achievement during the elementary school years (Belsky et al., 2007; NICHD ECCRN, 2002) and adolescence (Burchinal et al., 2014; Vandell et al., 2010), we found that the quality of child care was linked with higher math and language and literacy test scores at 54 months of age and these benefits remained largely the same throughout the next decade of life. Effect sizes for these associations were generally small and ranged from 5-12% of a standard deviation (see Table 5). And although there were periods during which the benefits of child care quality were not statistically significant (e.g., first grade for math and third grade for math and language/literacy),

there was *no* empirical evidence to suggest that these academic benefits changed over time (Wald χ2s = 0.01.-3.07, *ns*). That is, there was no indication of convergence or fadeout at least with respect to children's academic test scores, nor was there any evidence to suggest that these academic benefits grew as children progressed throughout the K-12 educational pipeline.

In contrast to child care quality from birth to 54 months of age, the quality of classrooms in elementary school was, with one exception, *not* associated with children's math or language and literacy performance through age 15, with effect sizes ranging from 0-4% of a *SD*, on average (see Table 5). In the one instance that a statistically significant difference did emerge, results suggested that higher quality elementary school classrooms resulted in higher math test scores at age 15, with an effect size of roughly 7% of a *SD*. It is also important to acknowledge that although the benefits of elementary school quality were more often than not statistically indistinguishable from zero, this does *not* mean that child care quality had a greater "effect" (see Gelman & Stern, 2006). In fact, the benefits of child care and elementary school quality were comparable in magnitude for children's math performance across all four waves of data collection (Wald χ2s = 0.28-2.37, *ns*) and for three of the four waves for language and literacy outcomes (Wald χ2s = 0.27-2.20, *ns*). We did, however, find that the benefits of child care quality were somewhat larger than that of classroom quality in elementary school for age 15 literacy performance (Wald $\chi 2 = 5.60$, $p < .05$), with an effect size difference of approximately 10% of a *SD*.

Multiplicative effects of education quality. We next estimated a series of models that considered the multiplicative benefits of education quality (i.e., child care X elementary school quality interactions). That is, we examined the degree to which the long-term benefits of child care were conditioned on children's later experiences in elementary school. Results from this

effort revealed evidence for moderation during fifth (for language and literacy) and ninth (for math and language and literacy) grade. When plotting these interactions, the benefits of child care quality were found to be considerably larger when children subsequently experienced high (versus low) quality classrooms during the elementary school year. More specifically, the average benefits of child care quality between birth and 54 months of age were 7-8% of a *SD*; however, for children who subsequently experienced high quality education in the elementary grades, these child care benefits grew larger by the end of fifth and ninth grade, with effect sizes of roughly 15-20% of a *SD* for language and literacy test scores, and roughly 18% of a *SD* by the end of ninth grade for math test scores. In contrast, for children who ended up in low quality classrooms during elementary school, child care quality showed *no* consistent nor significant association with children's academic test scores by age 15. It is of note, however, that for children's language and literacy (but not math) achievement, high elementary school quality was associated with amplified estimated long-term effects of both positive *and* negative early child care quality (i.e., the main effect of elementary school quality for literacy is considerably smaller than the interaction term, whereas the difference for math is smaller). That is, children who experienced low quality child care and high quality elementary schools demonstrated the least optimal language and literacy achievement in fifth and ninth grade.

Timing specific models. In addition to our omnibus models of elementary school quality, we also estimated a series of models that examined the periods during which children experienced higher quality classrooms during the elementary school years to determine whether the *timing* of high quality education mattered for the maintenance of child care benefits. As discussed earlier, these analyses are important as they can indicate whether the findings reported above are attributable to certain periods during which children experience higher (or lower)

quality education in elementary school. Similar to the prior set of main effects models reported above, child care quality and classroom quality are both in the models as main effect along with the covariates, omitting only interaction terms between our focal predictors.

Results from this effort revealed that classroom quality in first, third, and fifth grade did not have differential implications for children's academic achievement, with absolute effect sizes ranging from 0-8% of a *SD* (see Table 3). And there was no consistent evidence for moderation by any specific grade level; the raw effect sizes of the interaction terms were statistically indistinguishable from one another during fifth (Wald χ2s = 0.23-1.51 *ns*) and ninth grade (Wald χ2s = 0.02-3.37, *ns*). We also examined model fit across specifications using the Bayesian information criterion (BIC) statistic. Given two models fit on the same data, the model with the smaller BIC value is considered to be better fitting. According to these fit statistics, the omnibus (i.e., quality composited across grades) models were consistently superior to the timing specific models. Thus, when taken together, these results suggest that high quality education in first grade did not matter more (or less) for sustaining child care benefits as compared with high quality education in third and/or fifth grade; instead, what really appeared to matter was children's cumulative experience across the elementary school years.

Robustness checks. To ensure that our findings were robust, we estimated a series of supplemental models. First, we estimated a set of falsification tests that considered whether prospective indicators of classroom quality during the elementary school grades predicted and/or moderated the associations between early child care quality and children's math and language and literacy test scores at 54 months of age through third grade, logically impossible relations (i.e., the future predicting the past). Beyond being logically implausible, we estimated these models because they address the possibility that the results discussed above stem from the fact

that children in high quality child care end up in high quality elementary school classrooms (i.e., post-treatment bias). As can be seen in Supplemental Table 3, results from these analyses revealed that classroom quality during the elementary school years did *not* predict children's math or language and literacy performance at 54 months of age, nor did classroom quality in elementary school interact with child care quality for these outcomes. Put another way, children who experienced higher quality care were *not* more (or less) likely to end up in higher quality classrooms during the elementary grades, which is also demonstrated in our selection models that revealed that net of covariates that selected children into the different environments, child care quality was *not* linked with higher quality elementary school education (see Table 4). When looking at first and third grade, we also found no evidence to suggest that prospective indicators of classroom quality predicted children's past academic performance nor moderated the documented associations.

Next, families' household income was associated with both child care and elementary school quality in our multivariate prediction models and, thus, our findings may be confounded by families' socioeconomic status. We, thus, estimated additional models that included income x child care and income x elementary school quality indicators as covariates. Results from these analyses were the same as those discussed above (see Model 4 of Supplemental Table 2). Relatedly, considering that the earliest time point children's cognitive skills were assessed was at 15 months of age, our focal analyses had not controlled for children's cognitive abilities given that doing so would incorporate effects of child care environments. As a precaution, however, we estimated a series of models that examined child care quality at 15, 24, 36 months, and 54 months (i.e., excluding quality scores at 6), which allowed us to adjust for any sorting of children into higher quality programs as a result of higher skills (and child care quality during the earlier

waves). Results from bivariate analyses revealed that children with higher cognitive skills were more likely to attend higher quality child care ($r = 0.13$, $p < .001$) and elementary schools ($r = 0.06$, $p < .05$), but the magnitude of the correlations were small and these differences were no longer statistically significant after accounting for the other covariates. Moreover, and as can be seen in Model 5 of Supplemental Table 2, all findings were the same as those reported above when adjusting for children's earlier cognition. That is, net of children's cognitive skills at 15 months of age (and earlier child care quality at 6 months), children who experienced higher quality child care performed better in areas of math and language and literacy and these academic benefits grew larger in magnitude when coupled with higher quality elementary school education. Thus, these analyses indicate that our findings were *not* an artifact of higher functioning children attending higher quality programs.

As a means of addressing potential confounds, a number of advances have also been made in quasi-experimental methodology (e.g., propensity scores). Although traditional propensity score matching often requires the dichotomization (or categorization) of predictors, which results in the loss of valuable information, there are other types of propensity scores that allow us to preserve the continuous nature of our focal predictors. To this end, we implemented a covariate adjustment with the propensity score to further probe our primary model specification (for more information on this method, see: Austin, 2008, 2011). We included all of the variables outlined in Table 4 in addition to site fixed effects as predictors of child care and elementary school quality. And because income was associated with *both* child care and elementary school quality in our multivariate predictor models, we included SES interaction terms with all focal variables of interest. As part of this effort, we estimated two sets of models: (a) outcome models that excluded all covariates that were included in stage 1 of the propensity scores; and (b)

outcome models that included all of the covariates that were included in stage 1 of the propensity scores. It is also important to note that in the context of covariate adjustment using the propensity score, balance is assessed by examining the standardized difference between the predictors of interest, conditional on the propensity score (Austin, 2008). After balancing the child care and elementary school quality conditions (see Supplemental Table 4), we found that our propensity score models confirmed the conclusions discussed above (see Models 6 and 7 of Supplemental Table 2).

Finally, considering that many of the policy issues for quality child care and early childhood education more generally revolve around the experiences of 4-year-olds, we also estimated a series of models that focused on the preschool period specifically. As can be seen in Model 8 of Supplemental Table 2, and similar to Keys and colleagues (2013), we found no consistent main effects of child care quality at 54 months of age (net of earlier child care quality and cognitive functioning) and no consistent evidence for moderation, suggesting that our findings pertain more to children's *cumulative* child care and early education experiences during the early childhood years as opposed to the preschool year specifically.

Discussion

With the growing investments in children's education during the early childhood years as a means of strengthening children's long-term educational prospects (Duncan & Magnuson, 2013; Phillips et al., 2017; Yoshikawa et al., 2013), research and policy interest in understanding how to maintain the benefits of early care and education programs throughout middle childhood and adolescence has been on the rise. Although prior child care studies with the NICHD SECCYD (e.g., Belsky et al., 2007; Burchinal et al., 2014; NICHD ECCRN, 2002; Vandell et al., 2010, 2016) have greatly contributed to the discourse surrounding the ways in which early

childhood experiences drive success in life and school, to date, studies with these data have not fully considered the extent to which the long-term benefits of high quality child care vary as a function of children's subsequent educational experiences. In the current investigation, we sought to fill in this gap in knowledge.

Informed by the sustaining environments hypothesis (Bailey et al., 2016), the questions posed as part of this investigation, although correlational, point to the conditional nature of the long-term links between high quality child care and children's academic success in school. To that end, the results reported herein build on the existing literature that has considered structural processes of the school environment as sources of heterogeneity to focus on classroom processes as a source (e.g., Currie & Thomas, 2000; Lee & Loeb, 1995; Zhai et al., 2012). Results reveal that the predictive associations between high quality child care and academic skills at age 15 appear greater in magnitude when supported by high quality interactions between teachers and children in elementary school classrooms, but diminish otherwise. Although the associations between child care quality and children's math and language and literacy performance were generally small in magnitude on average, they grew two- to three-fold between 54 months and age 15 for children who subsequently experienced more stimulating and supporting teacher-student interactions across the elementary grades. In contrast, these associations with age 15 performance converged to close to zero for children who did not have these experiences in elementary school.

More specifically, there was little to no evidence to suggest that: (a) children's experiences during the elementary school years took precedence over their experiences in early childhood (e.g., Bradley et al., 1988); (b) children's experiences during early childhood mattered more than later classroom experiences; nor (c) the benefits of high quality education across birth

to and 54 months of age and the elementary school years were additive. Instead, the findings suggest that the magnitude of possible benefits of high quality early care and education across the first decade of life, on average are modest in size, and conditional depending on experience. Associations between quality of care and later academic functioning persisted over time as the benefits of those early experiences combined with subsequent high quality experiences in the classroom while, subsequent classroom experiences had very few benefits in the absence of prior high quality child care. Thus, although there is promise in investing in high quality child care as a means of improving children's longer-term educational success, the associations between quality of child care and outcomes upon which those investments would be predicated are stronger 6 to 10 years later for some children as compared with others, in part as a function of their experiences in elementary school classrooms.

Importantly, the results from the current investigation also indicate that the apparent persistence of child care "effects" was largely a function of the *cumulative* quality of experiences in classrooms across the elementary grades rather than experiences in any specific grade. At the same time, however, descriptive analyses reveal that children infrequently experience high quality education across the elementary school years, let alone across *both* early childhood and elementary school. In light of these descriptive patterns, careful attention must be paid to monitoring the educational opportunities provided to young children during the first decade of life. For these reasons, both researchers and stakeholders interested in maintaining the long-term educational gains that result from investments in early child care will need to consider children's subsequent educational experiences. That is, high quality early childhood programs alone are not as likely to meet the long-term needs of young children as when they combine with high quality classroom settings across their educational careers. And although there is good reason to think

cumulative effects are the way in which children accrue benefits from exposure over time, one alternative reason that could explain why we see cumulative rather than timing specific associations, is simply that the benefits of high quality at any given time—in child care or in elementary school classrooms—were small *and* inconsistent. If experiences in any specific elementary grade were more potent, we might then see more sizable associations with later performance in, for example, first grade.

It is also important to acknowledge that high quality classrooms during the elementary schools years not only amplified the long-term positive "effects" of child care, but also its negative "effects": children who experienced low quality child care and high quality elementary schools demonstrated the least optimal literacy (but not math) achievement in fifth and ninth grade. Put another way, higher quality elementary school education did *not* compensate for lower quality child care experiences. Although unexpected, prior studies have also found that children who have a misalignment in experiences across the pre-school to elementary transition do less well in school (e.g., Abry, Latham, & Bassok, 2015). As our results suggest, children who experience lower quality child care enter school with lower-level skills, and if they subsequently enter a higher quality elementary school environment, it is plausible that this type of setting might not meet their individual needs (e.g., Claessens et al., 2014; Lipsey, Farran, & Hofer, 2015) and result in poorer than expected performance. And because children are unlikely to experience continuous high quality education throughout their educational careers, what our results suggest is that we must more carefully consider how schools build on the diverse skills of individual children to optimize their long-term school success.

This possible disconnect between the preschool and elementary educational systems is particularly concerning for children's literacy performance as prior studies have found that

teachers interactions for the development of children's language and literacy skills is highly sensitive to children's individual needs (Connor, Morrison, & Katch, 2004; Foorman et al., 1998; Juel & Minden-Cupp, 2000). For example, children with stronger literacy skills have been found to make greater progress in classrooms that use a meaning-based approach than children with weaker literacy skills who demonstrate greater literacy development in classrooms with a code-based emphasis (Juel & Minden-Cupp, 2000). Thus, providing children with high quality education for a few years cannot overcome potential challenges that children and families experience, which is why aligning standards, curriculum, instruction, and assessments across educational systems is of utmost importance. Ultimately, these findings regarding the persistence of negative associations between low quality care and children's literacy development is just as worrisome as the convergence of academic test scores discussed in the literature.

Despite these contributions to the early childhood literature, the results of the current investigation should be interpreted in light of a few important limitations and future directions. First, because the design of the NICHD SECCYD is correlational in nature, caution is warranted when interpreting our findings and conclusions. Although it is true that analyses adjusted for a rich set of covariates—including children's earlier functioning and care quality—without an experimental design, causal claims cannot be made about the additive and multiplicative benefits of high quality education during the first decade of life. Nonetheless, long-term longitudinal studies with direct observations of children's educational experiences are rare and investigations that randomly assign children to different quality educational environments across the first decade of life are even more unlikely (for an exception see: Araujo, Carneiro, Cruz-Aguayo, & Schady, 2016). Thus, the results reported herein do provide correlational evidence to suggest that the benefits of high quality child care are more persistent if children subsequently experience

high quality education during the elementary school years. And even though correlation does not imply causation, a correlation is necessary for a causal effect.

Next, and perhaps just as importantly, despite the rich information on children and families available in the NICHD SECCYD, these data are not nationally representative. Considering that there have not been enough empirical studies examining the benefits of providing long-term investments in children's education, our ability to draw definitive conclusions is limited, and, therefore, replication with different samples is necessary before generalizations can be made (Duncan, Engel, Claessens, & Dowsett, 2014). Relatedly, the results from the current investigation should be interpreted in light of: (a) the counterfactual conditions at the time that these children experienced non-maternal child care (early to mid-nineties) and were enrolled in elementary school (early 2000s); and (b) the fact that these were largely middle class families. For instance, the last decade has seen a rise in quality improvement efforts and a growing number of children are spending more time outside of the home before kindergarten entry. Moreover, a growing number of services are being targeted at low-income populations, which is why continued effort is necessary to understand whether the patterns reported herein would replicate among children and families who experience greater hardship. In this context, some of our descriptive findings are even more concerning: if the children of middle class families receive inconsistent high quality education across the first decade of life, then how frequently do the most disadvantaged children across the country experience these opportunities? It is also important to acknowledge that the home environment has undergone changes and may be improving at differential rates across the income distribution (Bassok, Finch, Lee, Reardon, & Waldfogel, 2016; Kalil, Ziol-Guest, Ryan & Markowitz, 2016), which can have implications for child care effects. Thus, when studying child care experiences, it is important to keep in mind

cohort effects given the changing early life experiences of children.

Although our timing specific models revealed that the quality of classrooms in any grade level did not individually matter, this may stem from measurement error in the quality variables at one time point as compared with the average from our omnibus models. Irrespective of this possibility, however, our main effect estimates for classroom quality in elementary school were not too dissimilar from prior correlational and experimental estimates (0-10% of a *SD*; Araujo et al., 2016; Pointz, Rimm-Kaufman, Grimm, & Curby, 2009; Keys et al., 2013), which supports the general notion that elementary schools have small effects on children's educational trajectories. Finally, as part of the current study, we focused on just one potential explanation for the convergence of child care benefits, but there are other possibilities, such as the foot in the door hypothesis by Bailey and colleagues (2016), that requires continued empirical attention.

With these limitations and future directions in mind, this study used a prospective longitudinal dataset of American children and finds that the long-term academic benefits of high quality child care were conditioned by children's subsequent classroom experiences. Specifically, the small (but statistically significant) associations between high quality child care and children's academic performance were found to persist, and in fact accumulate over time, when coupled with higher quality classroom environments during the elementary school years, whereas in lower quality environments, there were no long-term benefits. Thus, despite the fact that researchers and policymakers who have sought to reduce deficits in children's long-term school success have been increasingly interested in the early childhood years, the results of this investigation highlight the importance of continued investments in children's education throughout the first decade of life, and not just in early childhood alone. What these results also make clear is that children's educational experiences across the first decade of life have little

benefit in the long-term in the absence of continuous high quality educational opportunities.

References

Abry, T., Latham, S., Bassok, D., & LoCasale-Crouch, J. (2015). Preschool and kindergarten teachers' beliefs about early school competencies: Misalignment matters for kindergarten adjustment. *Early Childhood Research Quarterly, 31*, 78-88. doi: 10.1016/j.ecresq.2015.01.001

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park, CA: Sage Publications.

Ansari, A., López, M. L., Manfra, L., Bleiker, C., Dinehart, L. H. B., Hartman, S. C. & Winsler, A. (2017). Differential third grade outcomes associated with attending publicly funded preschool programs for low-income, Latino children. *Child Development, 88,* 1743-1756. doi: 10.1111/cdev.12663

Ansari, A. (2017). The selection of preschool for immigrant and native-born Latino families in the U.S. *Early Childhood Research Quarterly, 41,* 149-160. doi: 10.1016/j.ecresq.2017.07.002

Ansari, A. (in press). The persistence of preschool effects from early childhood through adolescence. *Journal of Educational Psychology.*

Araujo, M. C., Carneiro, P., Cruz-Aguayo, Y., & Schady, N. (2016). Teacher quality and learning outcomes in kindergarten. *The Quarterly Journal of Economics, 131*, 1415-1453. doi: 10.1093/qje/qjw016

Austin, P. C. (2008e). Goodness-of-fit diagnostics for the propensity score model when estimating treatment effects using covariate adjustment with the propensity score. *Pharmacoepidemiology and Drug Safety, 17,* 1202–1217. doi:10.1002/pds.1673

Austin, P. C. (2011). An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behavioral Research, 46,* 399-424. doi:

10.1080/00273171.2011.568786

Bailey, D., Duncan, G. J., Odgers, C. L., & Yu, W. (2016). Persistence and fadeout in the impacts of child and adolescent interventions. *Journal of Research on Educational Effectiveness*, 1-33. doi:10.1080/19345747.2016.1232459

Bassok, D., Gibbs, C. R., & Latham, S. (2015). *Do the effects of early childhood interventions systematically fade? Exploring variation in the persistence of preschool effects*. EdPolicyWorks working paper. Retrieved from http://curry. virginia. edu/uploads/resourceLibrary/36_Preschool_Fade_Out. pdf.

Bassok, D., Finch, J., Lee, R., Reardon, S. F., & Waldfogel, J. (2016). Socioeconomic gaps in early childhood experiences, 1998 to 2010. *AERA Open, 2*, 1–22. doi: 10.1177/2332858416653924

Belsky, J., Vandell, D. L., Burchinal, M., Clarke-Stewart, K. A., McCartney, K., & Owen, M. T. (2007). Are There Long-Term Effects of Early Child Care? *Child Development*, *78*, 681–701. doi: 10.1111/j.1467-8624.2007.01021.x

Bradley, R. H., & Caldwell, B. M. (1984). The relation of infants' home environments to achievement test performance in first grade: A follow-up study. *Child Development*, 803-809. doi: 10.2307/1130131

Bradley, R. H., Caldwell, B. M., & Rock, S. L. (1988). Home environment and school performance: A ten-year follow-up and examination of three models of environmental action. *Child Development*, 852-867. doi: 10.2307/1130253

Burchinal, M. R., 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

Burchinal, M., Magnuson, K., Powell, D., & Hong, S. S. (2015). Early childcare and education. In R. M. Lerner (Ed.in Chief) & M. Bornstein & T. Leventhal (Vol. Eds.), *Handbook of child psychology and developmental science: Vol. 4. Ecological settings and processes* (7th ed., pp. 223–267). Hoboken, NJ: Wiley.

Claessens, A., Engel, M., & Curran, F. C. (2014). Academic content, student learning, and the persistence of preschool effects. *American Educational Research Journal, 51*, 403-434.

Coley, R. L., Votruba-Drzal, E., Collins, M. A., & Miller, P. (2014). Selection into early education and care settings: Differences by developmental period. *Early Childhood Research Quarterly, 29,* 319-332. doi:10.1016/j.ecresq.2014.03.006

Connor, C. M., Morrison, F. J., & Katch, L. E. (2004). Beyond the reading wars: Exploring the effect of child-instruction interactions on growth in early reading. *Scientific Studies of Reading, 8,* 305-336. doi: 10.1207/s1532799xssr0804_1

Costa, P. T., & McCrae, R. R. (1985). *The NEO personality inventory manual.* Odessa, FL: Psychological Assessment Resources.

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

Currie, J. & Thomas, D. (2000). School quality and the longer-term effects of Head Start. *Journal of Human Resources*. 35, 755-774. doi: 10.2307/146372

Dallaire, D. H., & Weinraub, M. (2007). Infant–mother attachment security and children's anxiety and aggression at first grade. *Journal of Applied Developmental Psychology, 28,* 477-492. doi: 10.1016/j.appdev.2007.06.005

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

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. (1981). *Peabody Picture Vocabulary Test–Revised.* Circle Pines, MN: American Guidance Service.

Elze, M. C., Gregson, J., Baber, U., Williamson, E., Sartori, S., Mehran, R., ... & Pocock, S. J. (2017). Comparison of propensity score methods and covariate adjustment: evaluation in 4 cardiovascular studies. *Journal of the American College of Cardiology*, *69*, 345-357.

Fuller, B., Bein, E., Bridges, M., Kim, Y., & Rabe-Hesketh, S. (2017). Do academic preschools yield stronger benefits? Cognitive emphasis, dosage, and early learning. *Journal of Applied Developmental Psychology*, *52,* 1-11. dfoi: 10.1016/j.appdev.2017.05.001

Gelman, A., & Stern, H. (2006). The difference between "significant" and "not significant" is not itself statistically significant. *The American Statistician*, *60*, 328-331.doi: 10.1198/000313006X152649

Gormley, W. T., Phillips, D. A., & Gayer, T. (2008). Preschool programs can boost school readiness. *Science, 320,* 1723–1724. doi:10.1126/science.1156019

Heckman, J. J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science*, *312*, 1900-1902. doi: 10.1126/science.1128898

Hill, C., Gormley, W., & Adelstein, S. (2015). Do the short-term effects of a strong preschool program persist?. *Early Childhood Research Quarterly, 32*, 60-79. doi:10.1016/j.ecresq.2014.12.005

Juel, C., & Minden-Cupp, C. (2000). Learning to read words: Linguistic units and instructional

strategies. *Reading Research Quarterly*, *35*, 458-492. doi: 10.1598/RRQ.35.4.2

Kalil, A., Ziol-Guest, K. M., Ryan, R. M., & Markowitz, A. J. (2016). Changes in income-based gaps in parent activities with young children from 1988 to 2012. *AERA Open*, *2*, 1-17. doi:10.1177/2332858416653732.

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

Lee, V. E., & Loeb, S. (1995). Where do Head Start attendees end up? One reason why preschool effects fade out. *Educational Evaluation and Policy Analysis*, *17*, 62-82. doi: 10.3102/01623737017001062

Lipsey, M. W., Farran, D.C., & Hofer, K.G. (2015). *A randomized control trial of a statewide voluntary prekindergarten program on children's skills and behaviors through third grade.* Nasville TN: Peabody Research Institute, Vanderbilt University.

Magnuson, K. A., Ruhm, C., & Waldfogel, J. (2007). The persistence of preschool effects: Do subsequent classroom experiences matter?. *Early Childhood Research Quarterly*, *22,* 18-38. doi: 10.1016/j.ecresq.2006.10.002

McCartney, K., Owen, M. T., Booth, C. L., Clarke- Stewart, A., & Vandell, D. L. (2004). Testing a maternal attachment model of behavior problems in early childhood. *Journal of Child Psychology and Psychiatry*, *45*, 765-778. doi: 10.1111/j.1469-7610.2004.00270.x

Medoff-Cooper, B., Carey, W. B., & McDevitt, S. (1993). The Early Infancy Temperament Questionnaire. *Journal of Developmental and Behavioral Pediatrics, 14,* 230–235.

Meyers, M. K., & Jordan, L. P. (2006). Choice and accommodation in parental child care

decisions. *Community Development, 3,* 53-70. doi:10.1080/15575330609490207

Muschkin, C. G., Ladd, H. F., & Dodge, K. A. (2015). Impact of North Carolina's early childhood initiatives on special education placements in third grade. *Educational Evaluation and Policy Analysis, 37,* 478-500. doi:10.3102/0162373714559096

Muthén, L. K., & Muthén, B. O. (1998-2013). *Mplus user's guide. Sixth edition.* Los Angeles, CA: Muthén & Muthén.

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 Institute of Child Health and Human Development, & Early Child Care Research Network. (2002). The relation of global first-grade classroom environment to structural classroom features and teacher and student behaviors. *The Elementary School Journal, 102,* 367-387.

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 and The Center for Child and Family Policy.

Pianta, R., Belsky, J., Houts, R., & Morrison, F. (2007). Opportunities to learn in America's elementary classrooms. *Science,* 315, 1795-1796. doi:10.1126/science.1139719

Pianta, R. C., Belsky, J., Vandergrift, N., Houts, R., & Morrison, F. J. (2008). Classroom effects on children's achievement trajectories in elementary school. *American Educational Research Journal, 45,* 365-397. doi: 10.3102/0002831207308230

Pianta, R. C., & Hamre, B. K. (2009). Conceptualization, measurement, and improvement of

classroom processes: Standardized observation can leverage capacity. *Educational Researcher, 38,* 109-119. doi: 10.3102/0013189X09332374

Ponitz, C. C., Rimm-Kaufman, S. E., Grimm, K. J., & Curby, T. W. (2009). Kindergarten classroom quality, behavioral engagement, and reading achievement. *School Psychology Review, 38,* 102-120.

Puma, M., Bell, S., Cook, R., Heid, C., Broene, P., Jenkins, F., ... & Downer, J. (2012). Third grade follow-up to the Head Start Impact Study: Final report. OPRE Report 2012–45, Office of Planning. *Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services, Washington, DC.*

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

Schaefer, E. S., & Edgerton, M. (1985). Parent and child correlates of parental modernity. *Parental belief systems: The psychological consequences for children,* 287-318.

Vandell, D. L., Belsky, J., Burchinal, M., Steinberg, L., & Vandergrift, N. (2010). Do effects of early child care extend to age 15 years? Results from the NICHD study of early child care and youth development. *Child Development, 81,* 737-756. doi:10.1111/j.1467-8624.2010.01431.x

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

Watamura, S. E., Phillips, D. A., Morrissey, T. W., McCartney, K., & Bub, K. (2011). Double jeopardy: Poorer social- emotional outcomes for children in the NICHD SECCYD

experiencing home and child- care environments that confer risk. *Child Development, 82,* 48-65. doi: 10.1111/j.1467-8624.2010.01540.x

Weiland, C., & Yoshikawa, H. (2013). Impacts of a prekindergarten program on children's mathematics, language, literacy, executive function, and emotional skills. *Child Development, 84,* 2112-2130. doi: 10.1111/cdev.12099

Woodcock, R. W., & Johnson, M.B. (1989). Woodcock-Johnson Psycho-Educational Battery-Revised. Allen, TX: DLM Teaching Resources.

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.

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

Table 1.
Descriptive statistics for covariates.

	Measurement period		
	Birth to 54 months of age	Elementary school years	Ninth grade
Child characteristics and experiences			
Male	0.52	--	--
White	0.81	--	--
Black	0.13	--	--
Other race	0.07	--	--
Birthweight (pounds)	7.71 (1.11)	--	--
Birth order	1.81 (0.92)	--	--
Temperament	3.18 (0.40)	--	--
Cognitive skills	108.55 (14.10)	--	--
Health	3.34 (0.74)	--	--
Secure attachment	0.61	--	--
Proportion in center care	0.20 (0.26)	--	--
Poor school quality	--	1.50 (0.31)	1.91 (0.49)
Family and household characteristics			
Maternal…			
Education	14.29 (2.51)	--	--
Age	28.19 (5.63)	--	--
Health	3.28 (0.73)	--	--
Locus of control	47.71 (7.54)	--	--
Maternity leave	0.47	--	--
Language skills	99.04 (18.32)	--	--
Psychological adjustment	59.06 (13.86)	--	--
Progressive child rearing beliefs	32.77 (3.53)	--	--
Traditional child rearing beliefs	59.98 (15.12)	--	--
Life stress	1.18 (0.75)	--	--
Positive feelings about pregnancy	4.09 (1.20)	--	--
Positive feelings about baby	1.89 (0.40)	--	--
Depression	9.33 (6.75)	8.71 (7.26)	10.41 (9.71)
Sensitivity	-0.02 (0.75)	-0.02 (0.85)	0.00 (1.00)
Parenting quality	-0.04 (0.85)	-0.00 (0.93)	0.01 (1.00)
Relationship quality with partner	5.72 (0.75)	3.87 (0.83)	3.80 (0.99)
Proportion of epochs in school	0.08 (0.18)	0.09 (0.21)	0.07 (0.26)
Proportion of epochs employed	0.66 (0.37)	0.76 (0.36)	0.80 (0.40)
Household income to needs	3.64 (2.86)	4.28 (3.47)	5.31 (5.82)
Proportion of epochs two parent household	0.84 (0.32)	0.81 (0.34)	0.77 (0.42)
Household size	4.12 (1.10)	4.35 (1.08)	4.19 (1.21)
Neighborhood disadvantage	10.27 (7.27)	7.37 (5.60)	7.04 (5.67)

Table 2.
Descriptive statistics for focal independent and dependent variables.

	Measurement period				
	Birth to 54 months of age [a]	First grade	Third grade	Fifth grade	Ninth grade
Language and literacy					
Picture vocabulary	459.57 (14.08)	483.99 (12.25)	496.94 (11.52)	505.82 (12.11)	518.64 (13.14)
Letter word identification	369.44 (21.43)	452.61 (24.02)	493.84 (18.73)	510.07 (17.56)	---
Word attack	---	473.95 (17.72)	494.58 (15.66)	---	---
Passage comprehension	---	---	495.31 (14.55)	505.16 (12.66)	520.43 (12.45)
Math					
Applied problems	424.77 (19.23)	470.10 (15.53)	497.30 (13.22)	509.82 (12.89)	524.54 (16.65)
Childcare quality	2.92 (0.42)	---	---	---	---
Classroom quality	---	5.10 (0.86)	5.25 (0.56)	5.23 (0.54)	---

Notes. [a] Children's academic achievement was measured at 54 months of age, whereas classroom quality was a composite measured from birth to 54 months.

Table 3.
Distribution of child care and elementary school quality.

	Elementary School Quality				
Child care quality	Low (<5.00)	Moderately Low (5.00-5.24)	Moderately High (5.25-5.49)	High (5.50-7.00)	Total
Low (1.00-2.74)	0.12	0.06	0.06	0.07	0.32
Moderately Low (2.75- 3.00)	0.08	0.04	0.05	0.07	0.25
Moderately High (3.01-3.29)	0.07	0.05	0.06	0.07	0.25
High (3.30-4.00)	0.04	0.03	0.05	0.06	0.18
Total	0.31	0.18	0.23	0.27	1.00

Notes. Proportions in rows and columns might not sum to total estimates due to rounding.

Table 4.
Bivariate and multivariate correlates of child care and elementary school quality.

	Child care quality		Elementary school quality	
Predictor	Bivariate	Multivariate	Bivariate	Multivariate
Child characteristics and experiences				
Male	-0.10 (0.06)	-0.06 (0.06)	-0.18 (0.06) **	-0.12 (0.06) *
Black	-0.61 (0.10) ***	-0.15 (0.12)	-0.70 (0.12) ***	-0.30 (0.14) *
Other race	0.11 (0.12)	0.21 (0.12)	0.05 (0.14)	0.08 (0.13)
Birthweight	0.00 (0.03)	-0.02 (0.03)	0.01 (0.03)	-0.02 (0.03)
Birth order	-0.08 (0.03) *	0.04 (0.04)	-0.03 (0.03)	0.03 (0.04)
Temperament	-0.04 (0.03)	-0.00 (0.03)	-0.08 (0.03) *	-0.03 (0.03)
Health	0.12 (0.03) ***	0.04 (0.03)	0.03 (0.03)	-0.03 (0.03)
Secure attachment	0.05 (0.06)	-0.01 (0.06)	0.14 (0.07) *	0.05 (0.06)
Proportion in center care	-0.16 (0.02) ***	-0.23 (0.03) ***	-0.03 (0.03)	-0.07 (0.04)
Child care quality	---	---	0.14 (0.03) ***	0.01 (0.03)
Family and household characteristics				
Maternal…				
Education	0.20 (0.03) ***	0.03 (0.04)	0.26 (0.03) ***	0.07 (0.04)
Age	0.15 (0.03) ***	-0.03 (0.04)	0.15 (0.03) ***	-0.10 (0.04) *
Health	0.11 (0.03) ***	0.06 (0.03) *	0.04 (0.03)	-0.03 (0.03)
Locus of control	-0.00 (0.03)	-0.01 (0.03)	0.01 (0.03)	0.06 (0.04)
Maternity leave	-0.06 (0.06)	-0.17 (0.07) *	0.10 (0.06)	-0.04 (0.07)
Language skills	0.17 (0.03) ***	0.01 (0.04)	0.25 (0.03) ***	0.03 (0.04)
Psychological adjustment	0.09 (0.03) **	-0.04 (0.04)	0.12 (0.03) ***	-0.01 (0.04)
Progressive childrearing beliefs	0.09 (0.03) **	0.01 (0.03)	0.11 (0.03) ***	0.04 (0.03)
Traditional childrearing beliefs	-0.16 (0.03) ***	0.03 (0.04)	-0.25 (0.03) ***	-0.05 (0.04)
Life stress	-0.04 (0.03)	0.03 (0.03)	-0.07 (0.03) *	0.00 (0.03)
Positive feelings about pregnancy	0.10 (0.03) ***	0.02 (0.03)	0.13 (0.03) ***	0.04 (0.03)
Positive feelings about baby	0.01 (0.03)	-0.01 (0.03)	0.04 (0.03)	-0.00 (0.04)
Depression	-0.11 (0.03) ***	0.02 (0.03)	-0.16 (0.03) ***	-0.02 (0.04)
Sensitivity	0.21 (0.03) ***	-0.01 (0.04)	0.31 (0.03) ***	0.12 (0.04) **
Parenting quality	0.36 (0.03) ***	0.32 (0.05) ***	0.29 (0.03) ***	0.05 (0.05)
Relationship quality with partner	0.08 (0.03) **	-0.01 (0.03)	0.13 (0.03) ***	0.06 (0.04)
Proportion of epochs in school	-0.04 (0.03)	-0.01 (0.03)	-0.00 (0.03)	0.06 (0.03) *
Proportion of epochs employed	-0.08 (0.04) *	-0.08 (0.03) *	0.04 (0.04)	0.03 (0.03)
Household income to need	0.19 (0.03) ***	0.10 (0.03) **	0.25 (0.03) ***	0.13 (0.03) ***
Proportion of epochs two parent household	0.22 (0.03) ***	0.06 (0.04)	0.21 (0.03) ***	0.01 (0.04)
Household size	-0.11 (0.03) ***	-0.09 (0.04) *	-0.02 (0.03)	0.05 (0.05)
Neighborhood disadvantage	-0.18 (0.03) ***	-0.02 (0.04)	-0.27 (0.04) ***	-0.04 (0.04)

Notes. All continuous variables have been standardized to have a mean of 0 and standard deviation of 1, and therefore, all coefficients correspond to effect sizes. Estimates in brackets correspond to standard errors. Bivariate correlates were estimated without FIML, whereas multivariate estimates addressed missing data. All estimates have been weighted by the proportion of epochs children attended non-maternal care and all variables were drawn from birth to 54 months of age.
*** $p < .001$. ** $p < .01$. * $p < .05$.

Table 5.
Multivariate results of children's academic achievement as a function of classroom quality during early (0-54 months) and middle (first to fifth grade) childhood

	Preschool		First grade		Third Grade		Fifth grade		Ninth Grade	
	Math	Language/ Literacy	Math	Language/ Literacy	Math	Language/ Literacy	Math	Language/ Literacy	Math	Language/ Literacy
Omnibus main effects model										
Child care quality	0.07 *	0.08 **	0.05	0.08 *	0.05	0.04	0.08 *	0.07 *	0.09 *	0.12 ***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.03)	(0.04)	(0.03)
Classroom quality during elementary school	--	--	-0.01	0.03	-0.02	0.04	0.01	0.01	0.07 *	0.02
	--	--	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Omnibus interaction model										
Child care quality X elementary school quality	--	--	0.03	0.03	0.00	0.05	0.05	0.07 *	0.09 *	0.07 **
	--	--	(0.03)	(0.03)	(0.04)	(0.03)	(0.04)	(0.03)	(0.03)	(0.03)
Timing specific main effect models										
Classroom quality in first grade	--	--	-0.01	0.03	-0.05	-0.00	-0.06	-0.03	-0.02	-0.01
	--	--	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Classroom quality in third grade	--	--	--	--	0.02	0.05	0.02	0.01	0.06	0.02
	--	--	--	--	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.03)
Classroom quality in fifth grade	--	--	--	--	--	--	0.06	0.03	0.08 *	0.03
	--	--	--	--	--	--	(0.03)	(0.03)	(0.04)	(0.03)
Timing specific interaction models										
Child care quality X first grade classroom quality	--	--	0.03	0.03	0.05	0.03	0.06	0.05	0.01	0.05
	--	--	(0.03)	(0.03)	(0.04)	(0.03)	(0.04)	(0.03)	(0.03)	(0.03)
Child care quality X third grade classroom quality	--	--	--	--	-0.05	0.02	0.03	0.01	0.10 *	0.03
	--	--	--	--	(0.04)	(0.04)	(0.04)	(0.03)	(0.04)	(0.03)
Child care quality X fifth grade classroom quality	--	--	--	--	--	--	-0.00	0.03	0.05	0.03
	--	--	--	--	--	--	(0.03)	(0.03)	(0.04)	(0.03)

Notes. All continuous variables have been standardized to have a mean of 0 and standard deviation of 1, and therefore, all coefficients correspond to effect sizes. Estimates in brackets correspond to standard errors. All estimates have been weighted by the proportion of epochs children attended non-maternal care.

*** $p < .001$. ** $p < .01$. * $p < .05$.