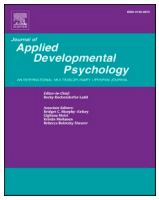
Journal of Applied Developmental Psychology 76 (2021) 101317

Contents lists available at ScienceDirect

Journal of Applied Developmental Psychology

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



Effects of North Carolina's pre-kindergarten program at the end of kindergarten: Contributions of school-wide quality

Robert C. Carr [a,*], Ellen S. Peisner-Feinberg [b], Rachel Kaplan [b], Irina L. Mokrova [a]

[a] Duke University, United States of America
[b] The University of North Carolina at Chapel Hill, United States of America

ARTICLE INFO

Keywords:
Pre-kindergarten
Elementary school
Language
Literacy
Mathematics
Working memory

ABSTRACT

The positive effects of pre-kindergarten (pre-K) programming may be enhanced in later grades for children who subsequently experience high-quality educational environments in elementary school. The current study tested this hypothesis in relation to the effects of North Carolina's NC Pre-K program on child outcomes at the end of kindergarten, including language, literacy, mathematics, and working memory. Measures of elementary school quality were examined as moderators of the NC Pre-K effects, including school-wide academic proficiency and school-wide growth in academic achievement. We found no reliable effects of NC Pre-K participation for children attending elementary schools with average levels of quality. However, the positive effects of NC Pre-K participation on language and working memory skills were evident for children attending elementary schools with higher levels of academic proficiency and academic growth, respectively. No evidence of moderation was found in relation to literacy and mathematics skills.

Introduction

Across the United States, publicly funded pre-kindergarten (pre-K) programs have been established by the federal government, 45 states, and numerous municipalities with the goal to promote children's school readiness (Friedman-Krauss et al., 2019). There is a wealth of evidence to indicate that participation in publicly funded pre-K programs can positively affect children's school readiness skills (e.g., Barnett et al., 2018; Lipsey, Farran, & Durkin, 2018; Puma et al., 2012). Although there is some evidence of long-term pre-K effects, studies have also found that the benefits of pre-K programming diminish during the early elementary grades as the test scores of pre-K and non-pre-K participants begin to converge as early as kindergarten (Phillips et al., 2017). This phenomenon—commonly referred to as fadeout or convergence—is not unique to public pre-K programs and has been found across studies of early childhood interventions more broadly (Bailey, Duncan, Odgers, & Yu, 2017; Bassok, Gibbs, & Latham, 2018). Moreover, the academic advantages gained from attending a pre-K program may be expected to diminish over time unless children continue to experience high-quality educational environments in elementary school (Bailey et al., 2017; Phillips et al., 2017). In the current study, we examined the effects of North Carolina's statewide pre-K program (NC Pre-K) on children's language, literacy, mathematics, and working memory skills at the end of kindergarten—one year after children had finished the NC Pre-K program. We also examined whether the effects of NC Pre-K varied across levels of elementary school quality.

The current study considered the effects of NC Pre-K, which is an established, high-quality pre-K program with a history of research documenting its effectiveness. We focused on a single state-funded pre-K program, given that programs have been shown to differ in their implementation as well as in their effectiveness (Barnett et al., 2018; Phillips et al., 2017). Additionally, to draw implications for educational policymaking, the current study focused on two policy-relevant indicators of elementary school quality—school-wide academic proficiency and school-wide academic growth—which are common targets of state-wide administrative policy and school accountability/reform initiatives in NC and nationally.

North Carolina's Pre-K program

Established in 2001, NC Pre-K is a state-funded educational program for eligible 4-year-olds, designed to enhance their school readiness

* Corresponding author at: Duke University, 2024 West Main Street, Duke Box 90539, Durham, NC 27708-0539, United States of America.
E-mail address: robert.carr@duke.edu (R.C. Carr).

https://doi.org/10.1016/j.appdev.2021.101317
Received 19 May 2020; Received in revised form 1 August 2021; Accepted 4 August 2021
Available online 20 August 2021

skills.[1] While NC was among the latter half of states to adopt a public pre-K program (Cohen-Vogel, Sadler, Little, Merrill, & Curran, 2020), NC Pre-K is now considered to be an established program, having operated at scale for nearly 20 years. During the 2014–15 program year—the year in which the current study's participants were enrolled in the NC Pre-K program—22% of the state's 4-year-olds participated in the program ($N = 26{,}851$; Barnett et al., 2015). According to program guidelines, children were eligible to enroll if they had turned four years old by August 31, 2014 and had a gross family income at or below 75% of the state median income (NC Division of Child Development and Early Education [DCDEE], 2015). Within a local program, up to 20% of age-eligible children with higher family incomes could have enrolled in the program if the child had at least one of the following additional factors: limited English proficiency, identified disability, chronic health condition, educational need as indicated by results from developmental screening, or a parent serving in the military.

NC Pre-K provides funding to serve eligible children in a variety of classroom-based educational settings, including public schools, Head Start, and private child care centers (both for-profit and nonprofit). Local sites were expected to meet a variety of program standards designed to ensure a high-quality, classroom-based educational experience for children and uniformity in the program across the state, to the extent possible (DCDEE, 2015). Specifically, the sites operated on a school day and school year calendar basis for 6–1/2 h/day and 180 days/year. Class sizes were restricted to 18 children with a lead and assistant teacher, with adult:child ratios of 1:9. Lead teachers were required to hold or be working toward a NC Birth through Kindergarten (B-K) teaching license, or the equivalent, and assistant teachers were required to hold or be working toward an Associate Degree in early childhood education or child development or a Child Development Associate credential. Classroom activities and instruction were based on the state's early learning standards (NC Foundations Task Force, 2013) and an approved curriculum. Classroom staff were expected to conduct developmental screenings and ongoing assessments to gather information on individual children's skill development in order to inform instruction.

Program standards also required regular evaluation and monitoring of all sites to ensure that high-quality services were being implemented and maintained. For example, individual NC Pre-K classrooms were required to demonstrate a minimum score of 5.0 on the Early Childhood Environment Rating Scale–Revised (Harms, Clifford, & Cryer, 1998), and improvement plans were developed for classrooms that scored below this threshold (DCDEE, 2015). The results of two state-wide studies of NC Pre-K providers document that program quality generally falls within the medium-quality to high-quality range across numerous dimensions (Peisner-Feinberg et al., 2014; Peisner-Feinberg, Schaaf, Hildebrandt, & LaForett, 2013). Similar to most other pre-K samples, direct observations of NC Pre-K classroom quality documented higher scores for dimensions of quality related to the global classroom environment, the language and literacy environment, classroom management and organization, and the emotional sensitivity/supportiveness of teacher-child interactions (Peisner-Feinberg et al., 2013, 2014). Alternatively, dimensions of quality related to instructional support tend to be lower in NC Pre-K classrooms—findings that are consistent with research on pre-K programming more broadly (Pianta, Downer, & Hamre, 2016). Moreover, the NC Pre-K program has historically met the quality standard benchmarks established by the National Institute of Early Education Research; meeting all ten benchmarks during the 2014–15 program year in which the current study's participants were enrolled in NC Pre-K (Barnett et al., 2015).

[1] In 2011, the NC General Assembly transferred the existing program from the NC Department of Public Instruction (DPI) to the Division of Child Development and Early Education (DCDEE) in the Department of Health and Human Services (DHHS) and renamed it from the More at Four Pre-K Program to the NC Pre-K Program.

NC Pre-K program effectiveness

The effectiveness of the NC Pre-K program has been documented across numerous studies that examined children's academic skills. Although a number of well-designed studies found positive effects over time, there is some mixed evidence with regard to specific types of outcomes. Two program evaluation studies conducted by Peisner-Feinberg and colleagues examined the effects of NC Pre-K on children's school readiness skills (Peisner-Feinberg & Schaaf, 2011; Peisner-Feinberg, Zadrozny, Kuhn, & Van Manen, 2019). One statewide study utilized a regression discontinuity design and found positive effects of NC Pre-K participation on children's literacy (Cohen's $d = 0.56$–1.16) and mathematics ($d = 0.34$–0.81), but no reliable effect on language skills (Peisner-Feinberg & Schaaf, 2011). In a small-scale randomized control trial (RCT) of children in two of NC's 100 counties, positive NC Pre-K effects were found on children's end-of-pre-K language ($d = 0.22$) and literacy skills ($d = 0.27$), but no reliable effects on mathematics or executive functioning skills were found (Peisner-Feinberg et al., 2019). While positive effects on children's school readiness literacy skills were consistent across these two studies conducted nearly a decade apart, positive effects on mathematics and language skills varied across the two studies.

There is additional evidence to indicate that the positive effects of NC Pre-K are evident in elementary school and later grades. A statewide study by Peisner-Feinberg, Mokrova, and Anderson (2017) examined the effects of NC Pre-K participation on child outcomes at the end of kindergarten based on propensity score matching and found evidence of favorable program effects on measures of language ($d = 0.15$), mathematics ($d = 0.16$–0.22), and working memory ($d = 0.17$) but no reliable effect on literacy skills. Alternatively, the small-scale RCT study found positive NC Pre-K effects on children's end-of-pre-K language and literacy skills, but these effects were no longer evident at the end of kindergarten (Peisner-Feinberg, Kuhn, Zadrozny, Foster, & Burchinal, 2020). Separate statewide studies documented evidence of NC Pre-K effects on reading and mathematics achievement in third grade, along with effects on measures of school progress (i.e., reductions in special education placements and grade retention), both for the comparison of NC Pre-K participants and non-participants (Peisner-Feinberg & Schaaf, 2010) and for the overall population of children in NC (Dodge, Bai, Ladd, & Muschkin, 2016; Ladd, Muschkin, & Dodge, 2014; Muschkin, Ladd, & Dodge, 2015). The latter findings were drawn from a series of studies that exploited variation in NC Pre-K program funding across counties and years to examine the population-wide effects of increased NC Pre-K funding on the outcomes of all children, including both former NC Pre-K participants as well as non-participants—referred to as *spillover* effects (Bai, Ladd, Muschkin, & Dodge, 2020; Dodge et al., 2016; Ladd et al., 2014; Muschkin et al., 2015). These studies found that increased funding for NC Pre-K produced positive effects on all children's third grade reading and mathematics achievement ($d = 0.056$–0.064) along with reductions in special education placements and grade retention, which continued into eighth grade. In sum, positive NC Pre-K program effects on children's academic skill development have been observed across a wide range of child outcomes and grade levels, and across numerous study designs.

Fadeout and convergence

Although the aforementioned studies provide some evidence to indicate that NC Pre-K has favorable short-term and long-term effects, evaluation studies of three other large-scale pre-K programs documented evidence of pre-K effects diminishing over time. The terms fadeout and convergence describe how any initial difference in the test scores of pre-K and non-pre-K participants proceed to diminish over time through a process of stunted progress among former pre-K participants

and/or accelerated growth among the comparison group. For example, random assignment studies of Tennessee's statewide pre-K program (Lipsey et al., 2018) and the federal Head Start program (Puma et al., 2012) found evidence of test score convergence as early as kindergarten, with limited evidence of persistent program effects into third grade in the Head Start study. Quasi-experimental studies of Tulsa, Oklahoma's universal pre-K program also demonstrated mixed findings, with evidence of convergence during elementary school as well as positive effects in middle school (Gormley, Phillips, & Anderson, 2018; Hill, Gormley, & Adelstein, 2015). A meta-analysis of ECE program effect studies across five decades found that convergence in children's cognitive test score outcomes was most pronounced during kindergarten (Li et al., 2016), pointing to kindergarten as a particularly salient period to examine factors that may lead to convergence—the focus of the current study.

Many scholars have speculated that the long-term effects of early childhood interventions will depend on high-quality educational inputs in elementary school (Bailey et al., 2017; Cunha & Heckman, 2007; Phillips et al., 2017). For example, the concept of dynamic complementarity suggests that early investments in children's skill development (e.g., high-quality pre-K) are bolstered by subsequent investments in children's skill development (e.g., high-quality schooling) (Cunha & Heckman, 2007). Moreover, early investments will *not* be productive in the long-term unless they are followed up by subsequent investments. The sustaining environments hypothesis specifically identifies high-quality schooling as a subsequent investment needed to bolster the effects of early childhood intervention (Bailey et al., 2017). To address this issue, education stakeholders have become increasingly interested in the design of initiatives to align educational policies and practices across preschool through elementary school (e.g., P-3 or even P-12) and ensure continuity in high-quality learning experiences for children (e.g., Stipek, Clements, Coburn, Franke, & Farran, 2017). As education stakeholders work to ensure that public investments in pre-K programming carry over into later school success, they will benefit from understanding whether subsequent school quality contributes to enhancing the effects of pre-K programming. In the current study, we sought to investigate if two dimensions of subsequent school quality enhance the effects of NC Pre-K.

School quality

There is a wealth of evidence to suggest that the quality of educational opportunity in schools is related to children's learning and developmental outcomes (Eccles & Roeser, 2010). Beginning with the enactment of the federal No Child Left Behind Act, all states were mandated to assess and report on measures of school quality for accountability and quality monitoring purposes. In the current study, we examined two policy-relevant measures of school quality based on aggregate student test scores in reading and mathematics: (1) school-wide academic proficiency as indexed by the percent of students in the school who score in the "proficient" range at a single point in time, and (2) school-wide growth in academic achievement as indexed by the rate at which students grow or gain in their academic test scores across grades.

Using data from 11,000 school districts across the United States, Reardon (2019) found a small correlation between academic achievement in third grade and academic growth from third to eighth grade ($r = -0.13$), suggesting these two measures index distinct dimensions of school quality. This study also found that third grade academic achievement showed a strong positive correlation with district-wide socioeconomic status ($r = 0.68$), while academic growth showed a much smaller correlation ($r = 0.32$; Reardon, 2019). Based on these findings, Reardon (2019) argued that academic growth may better reflect the quality of educational opportunity that schools uniquely provide to students. Nonetheless, both measures—to varying degrees—will reflect student's access to educational opportunities (or lack thereof) derived from their family backgrounds, their broader community resources, as well as their schools. Moreover, high-stakes accountability and quality monitoring systems motivate education agencies to improve the academic proficiency and academic growth of students from all family backgrounds. In the current study, we considered measures of school-wide academic proficiency and academic growth because they index two, seemingly distinct and policy-relevant dimensions of the quality of educational opportunity in schools. There are myriad ways in which the academic performance of peers in a school can influence the academic achievement of individual students, which are discussed next.

Mechanisms of school quality: The academic peer environment

A robust body of research has found that an individual's achievement is enhanced in the context of a school environment with higher-performing peers or diminished in an environment with lower-performing peers (e.g., Hanushek, Kain, Markman, & Rivkin, 2003). Social norms theory suggests that an individual member of a group is likely to become engaged with those behaviors and attitudes that are perceived to be acceptable by other group members (e.g., Cialdini & Trost, 1998). In the context of a high-performing peer environment, a student may perceive the peer group to be academically engaged and may be more likely to become academically engaged themselves, while the opposite may be evident in the context of a low-performing peer environment (Lynch, Lerner, & Leventhal, 2013). Peers often interact with one another during classroom activities (both one-on-one and in groups) and can serve as co-learners—actively helping each other to learn academic content by interpreting and clarifying concepts and tasks, and modeling positive academic behavior (Eccles & Roeser, 2010). Alternatively, a low-performing peer environment may be characterized by frequent classroom disruptions and challenging behavior, which may hinder learning opportunities for individual students (Lavy, Paserman, & Schlosser, 2012).

While the academic peer environment of a school may be directly related to the academic achievement of individual students, it may also be indirectly related through impacts on the instruction provided by teachers. For example, teachers in schools with more high-performing students may feel more motivated, maintain greater expectations for student learning, teach at a faster pace, and/or cover more advanced academic content when they perceive students as possessing more advanced abilities (Henry & Rickman, 2007). Additionally, classroom disruptions and challenging behavior, which may be characteristic of schools with more low-performing students, may hinder opportunities for teachers to focus their time and attention on academic instruction (Henry & Rickman, 2007). Regardless of the specific mechanism by which peers can promote or hinder the academic achievement of an individual student, there is consensus that the academic performance of a student's peers provides an influential source of socialization for academic achievement (Eccles & Roeser, 2010).

Correlates of school quality

The current study also considered two school-level factors that may be related to measures of school-wide academic proficiency and academic growth: school-level poverty and school size. In relation to school-level poverty, there is consistent evidence that the socioeconomic makeup of students in a school continues to be a strong predictor of the average academic achievement level of the school despite decades of school reform efforts—with schools serving larger proportions of low-income students reporting lower levels of school-wide academic proficiency (Battistich, Solomon, Kim, Watson, & Schaps, 1995; Reardon, 2019; Vanderhaar, Muñoz, & Rodosky, 2006). Not surprisingly, predominantly low-income schools are more likely to be located in communities with higher rates of environmental stressors such as disorganization and violence that can impede children's learning and development (McCoy, Roy, & Sirkman, 2013; Raver, McCoy,

Lowenstein, & Pess, 2013). Moreover, this environmental stress may be reflected within the schools themselves, as schools characterized by high levels of student poverty have been found to have larger class sizes and fewer classroom aides (National Institute of Child Health and Human Development Early Child Care Research Network, 2004). Community-level and school-level stressors may also produce challenging work-place conditions for teachers that can lead to reductions in school quality, as high-poverty schools often struggle to attract and retain qualified teachers (Ronfeldt, Loeb, & Wyckoff, 2013). Therefore, in the current study, we controlled for school-wide poverty as measured by the percent of economically disadvantaged students in the school.

We also controlled for school size in our study, as measured by the total number of students enrolled in the school. Decades of research into the effects of elementary school size has documented some evidence to indicate that attendance in a larger school has a negative impact on student achievement (Gershenson & Langbein, 2015; Kuziemko, 2006). For example, a study by Gershenson and Langbein (2015) used school-level administrative data from NC to examine the effects of school size in relation to student achievement in fourth and fifth grade. Analyses of the full sample revealed no reliable effect of school size in relation to academic achievement, but found that two subgroups—socioeconomically disadvantaged students and students with learning disabilities—performed worse in larger schools.

Early childhood education effects and elementary school quality

Numerous studies have investigated measures of school quality in elementary school in relation to the long-term effects of early childhood education (ECE) programming. While most studies focused on school-wide academic proficiency, a handful also considered school-wide academic growth. While some studies specifically considered publicly funded ECE programs (e.g., Curenton, Dong, & Shen, 2015; Jenkins et al., 2018; Pearman et al., 2020; Unterman & Weiland, 2020), other studies have combined publicly funded and community based ECE programs (Ansari & Pianta, 2018) or examined curricular interventions based in Head Start classroom settings (Bierman et al., 2014; Zhai, Raver, & Jones, 2012). Findings from several studies provide evidence to indicate that ECE program effects were enhanced under higher-quality elementary school contexts, while other studies found contrasting evidence.

Evidence of dynamic complementarity was provided in several studies. For example, Zhai et al. (2012) found intervention effects of the Chicago School Readiness Project at the end of kindergarten for children who were subsequently enrolled in elementary schools categorized as high in school-wide academic proficiency (i.e., 0.5 standard deviations above the average), but no reliable effects were found for children in schools categorized as low proficiency (i.e., 0.5 standard deviations below the average). Similarly, a study by Ansari and Pianta (2018) found a reliable effect of ECE participation through fifth grade for children in moderate-quality and high-quality elementary schools, but not low-quality elementary schools. In this study, school quality was represented by an aggregate index of 20 indicators—including school-wide academic proficiency—measured across four grades and averaged over time. Support for dynamic complementarity was also provided in a study of the Boston Public Schools Pre-K program's effects on children's third-grade outcomes—suggesting that higher-levels of school-wide academic proficiency enhanced the program's effect on third-grade mathematics scores, but not English language arts scores or school progress measures, while measures of school-wide academic growth did not moderate the program's effect (Unterman & Weiland, 2020). Finally, reanalysis of the random assignment study in Tennessee found no reliable evidence to indicate that a measure of school-wide academic growth by itself moderated the effects of Tennessee's Voluntary Pre-K program on third-grade outcomes, but the Tennessee Pre-K effect was found to persist under the combination of high levels of school-wide academic growth and highly effective teachers in elementary school (Pearman et al., 2020).

In contrast, a study by Bierman et al. (2014) found that the effect of the Head Start REDI intervention on children's social-behavioral skills was only evident through the end of kindergarten for children who subsequently enrolled in low-proficient elementary schools, as measured by the level of school-wide academic proficiency of all third-grade students in the school. Additionally, a study by Curenton et al. (2015) found that the positive effects of pre-K participation on children's fifth-grade math and reading achievement was weakened when the level of school-wide academic proficiency of fifth-grade students was considered as a mediating variable.

Studies also have found null effects, including a reanalysis of the Head Start Impact Study by Jenkins et al. (2018), which found no reliable evidence to indicate that measures of school-wide academic proficiency moderated the effects of Head Start participation on kindergarten outcomes. Moreover, findings from a meta-analysis of studies—including many of the aforementioned studies—provided no clear consensus to identify subsequent school contexts under which the effects of ECE programming are enhanced (Bailey, Jenkins, & Alvarez-Vargas, 2020). Many scholars have speculated about the role of higher-quality elementary school context in relation to ECE program effects, but the available evidence is mixed—indicating that ECE effects were more likely to be enhanced in both higher-quality and lower-quality school environments, or not at all.

The current study

The current study examined the effects of children's participation in the NC Pre-K program on child outcomes at the end of kindergarten—one year after children had finished participating in the program. We simultaneously examined if characteristics of elementary school-wide quality enhanced the effects of NC Pre-K. While prior studies documented positive effects of NC Pre-K at school entry and into later grades, the current study was the first to examine the role of elementary school-wide quality during kindergarten in relation to NC Pre-K effects at the end of kindergarten. Prior research indicates that the kindergarten year is a period of time in which early childhood program effects are most likely to diminish (Li et al., 2016). The primary research questions of the current study were as follows: (1) To what extent does NC Pre-K participation affect children's end-of-kindergarten skills at the average level of elementary school quality? (2) To what extent do the effects of NC Pre-K participation vary across higher and lower levels of elementary school quality? Consistent with the concept of dynamic complementarity and the sustaining environments hypothesis (Bailey et al., 2017; Cunha & Heckman, 2007), we hypothesized that the favorable effects of NC Pre-K participation would only be evident at the end of kindergarten for students in high-quality elementary schools.

Methods

Sample

We undertook secondary analyses of data from a state-wide evaluation of the NC Pre-K program that was described in a previous report by Peisner-Feinberg et al. (2017). Two groups of children were compared—those who participated in NC Pre-K (treatment) and those who did not participate in NC Pre-K (comparison) during the 2014–2015 program year and then enrolled in kindergarten during the 2015–2016 school year. Our study utilized data collected from a survey of children's parents, direct assessments of children's developmental skills at the end of kindergarten, and state-wide administrative data collected by the NC Department of Public Instruction (DPI).

Parent surveys were collected in the fall of kindergarten and provided demographic information about children and families. Based on information for children whose families returned the survey on time, an initial pool of 1355 kindergarten children was established for the study.

Children were excluded from this initial pool if they met any of the following exclusion criteria: had been retained in grade; had a birthdate outside of the treatment window (before 9/1/09 or after 8/31/10); did not meet NC Pre-K income eligibility criteria; had an IEP; attended an NC Pre-K site but were not funded through NC Pre-K (to avoid confounds between the treatment and comparison groups)[2]; or had incomplete survey data. This procedure resulted in an input sample of 823 children, including treatment ($n = 432$) and comparison ($n = 391$) cases. Based on this input sample, propensity score matching was used to select a matched sample of 532 children with similar characteristics across the treatment ($n = 266$) and comparison groups ($n = 266$; see *Propensity Score Matching* section below).

A battery of child assessments was administered directly to the matched sample of treatment and comparison group children in the late spring of their kindergarten school year (4/18/16–5/27/16). Children were assessed on a range of measures appropriate for kindergartners across four primary areas—language, literacy, math, and working memory. Individual child assessments were conducted in the schools by trained data collectors. Information about each of the 33 elementary schools that children attended during their kindergarten year (i.e., 2015–2016) was obtained from a statewide administrative database housed by the NC DPI, Accountability and Services Division (www.ncpublicschools.org).

Sampling procedure

A complex sampling procedure was used to select participants for the current study. A total of 24 counties in NC were identified as study sites in order to represent regions of the state (North, South, East, West), proportion of children in the NC Pre-K program (low, medium, high by terciles), and total number of children enrolled in NC Pre-K. A stratified random sample of 24 counties was selected from the 100 counties in NC based on region and proportion, weighted by the number enrolled. Two counties were randomly selected within each of the 12 regions by proportion cells. A selected county was replaced with the next county on the randomly generated list within a given cell in cases where the school district(s) declined to participate. A total of 33 school districts within 31 counties were initially contacted for participation in the study, with a final sample of 24 districts in 24 counties (77%). Within those 24 districts, the principals of 44 schools were contacted about the study and 35 agreed for their schools to participate (80%). Information describing the study, parent permission forms, and the parent survey forms were sent home to all 2843 parents within these classrooms; 1524 returned the survey and gave permission for their children to participate in the study (54%).

Measures

NC Pre-K participation

Our treatment variable of interest was defined by whether children participated in the NC Pre-K program during the year prior to kindergarten. Information from the parent survey was used to identify children as either an NC Pre-K participant or a comparison case (did not participate in NC Pre-K) based on a series of survey questions about pre-K participation (e.g., "Was your child enrolled in the NC Pre-K Program last year [2014–2015] as a 4-year-old [was their pre-K education funded by NC Pre-K last year]?"). The accuracy of this information was then validated against the NC Pre-K state administrative database.

Pre-K classroom membership

Parents were asked to identify the name of the pre-K site and teacher of the classroom in which their child was enrolled in during the pre-K year—the year prior to the study year—and this information was used to categorize children's pre-K classroom membership. As defined by the exclusion criteria detailed in our *Sample* section above, children who attended an NC Pre-K site but were not funded through NC Pre-K (based on NC Pre-K administrative records) were excluded from the sample to avoid confounds between the treatment and comparison groups. Therefore, no treatment and comparison group children were enrolled in the same pre-K classroom.

Child outcomes

Children's language, literacy, mathematics, and working memory skills were assessed at the end of kindergarten.

Language. Children's receptive and expressive language skills were assessed with the Picture Vocabulary subtest of the Woodcock-Johnson (WJ) III Tests of Achievement (Woodcock, McGrew, & Mather, 2001). The WJ III is a norm-referenced measure, and standard scores were used in our analyses. The measure developers report internal consistency reliability as $r = 0.70$–0.76 for children in the 5- to 6-year-old age range (McGrew & Woodcock, 2001).

Literacy. Literacy skills were assessed with the WJ III Brief Reading Composite, which was comprised of two subtests from the WJ III Tests of Achievement: Letter-Word Identification and Passage Comprehension. The WJ III Letter-Word Identification subtest measured basic pre-reading and reading skills, including letter and word recognition and identification skills. The WJ III Passage Comprehension subtest measured symbolic learning and basic comprehension skills. Standard scores were used in our analyses of these measures. The measure developers report internal consistency reliability as $r = 0.98$–0.99 for the WJ III Letter-Word Identification subtest and $r = 0.96$ for the WJ III Passage Comprehension subtest for children in the 5- to 6-year-old age range (McGrew & Woodcock, 2001).

Mathematics. Mathematics skills were assessed with the WJ III Brief Math Composite, which was comprised of two measures from the WJ III Tests of Achievement: Applied Problems and Calculation. The WJ III Applied Problems subtest measured math problem-solving skills including simple comparisons, counting, addition, and subtraction. The WJ III Calculation subtest measured basic mathematical computation skills. Standard scores were used in our analyses of these measures. The measure developers report internal consistency reliability as $r = 0.96$–0.97 for the WJ III Calculation subtest and $r = 0.88$–0.92 for the WJ III Applied Problems subtest for children in the 5- to 6-year-old age range (McGrew & Woodcock, 2001).

Short-term working memory. Short-term working memory was assessed using the Forward Digit Span and Backward Digit Span measures, which assess the phonological loop and central executive functioning components of working memory, respectively (Gathercole & Pickering, 2000). During these tests, children were asked to recall a sequence of digits that were spoken out loud by the assessor in normal order (i.e., forward) and reverse order (i.e., backward). For our analyses, we calculated a composite measure of short-term working memory as the average of the standardized scores from the Forward and Backward Digit Span assessments. We utilized a composite measure of children's short-term working memory skills based on evidence from psychometric studies suggesting that a single factor can adequately characterize children's underlying latent ability with respect to performance on these tasks during early childhood (Fuhs & Day, 2011; Willoughby, Wirth, & Blair, 2011). This is also consistent with prior studies of children participating in public pre-K programs which used composite scores based on the specific measures used in the current study (Lipsey et al., 2017; Weiland & Yoshikawa, 2014). Test-retest reliability has been reported as 0.72 for these combined tests for children in the 6- to 7-year-old age range on

[2] The median proportion of NC Pre-K funded children per NC Pre-K classroom was >90%.

related versions of these measures (Wechsler, 2014).

Elementary school-wide quality

Two measures of elementary school-wide quality were considered in the current study, including academic proficiency and academic growth, which were provided by the NC DPI.

Academic proficiency. School-wide academic proficiency was measured by the percent of third-grade students in the school who scored in the "proficient" range on the end-of-grade (EOG) reading and mathematics tests (i.e., the average of reading proficiency and mathematics proficiency). We considered school-wide academic proficiency at third grade because it was the lowest grade level at which standardized achievement tests were administered in NC.

Academic growth. School-wide growth in academic achievement was measured using the EVAAS granular growth score provided by the NC DPI (2016). School-wide academic growth was calculated based on the math and reading EOG test scores of all available students in the school from third grade through the school-maximum grade.

Elementary school covariates

Our analyses also controlled for the level of *school-wide poverty* (i.e., the percent of economically disadvantaged students) and *school size* (i.e., the total number of students enrolled in the school) to account for these potential sources of bias. Data were provided by the NC DPI.

Child and family background covariates

Measures of child and family background characteristics were included as covariates in the propensity score matching and regression analyses. Covariates included gender (Female = 0, Male = 1), Hispanic/Latino ethnicity (Non-Latino = 0, Latino = 1), race (Black/African-American, Native American, Asian, with White as a reference cell), chronic health condition (No = 0, Yes = 1), child speaks English at home (No = 0, Yes = 1), English is child's first language (No = 0, Yes = 1), frequency of English spoken at home by primary caregiver (Never = 0, Rarely = 1, Sometime = 2, Frequently = 3, Always = 4), parent in the military (No = 0, Yes = 1), primary caregiver education (8th grade or less = 1, Some High School = 2, High School Diploma/GED = 3, AA/AS degree = 4, BA/BS degree = 5, MA/MS degree = 6, Doctorate degree = 7, N/A = 8), family size, number of adults in the home, number of children in the home, family income ($26,069 or less = 1, $26,070–$34,090 = 2, …, $73,670–$75,200 = 12, $75,201 or more = 13), and whether the parent survey was completed in English (No = 0, Yes = 1).

Propensity score matching

Propensity score matching (PSM) was used to select a matched sample of children across the treatment and comparison groups. Based on the input sample of 823 children in the treatment (n = 432) and comparison (n = 391) groups, logistic regression was used to calculate the propensity score, which represents the likelihood of treatment as a function of the aforementioned child and family background covariates (with county included as a fixed effect). Next, propensity score matching was conducted using the MatchIt function in R (Ho, Imai, King, & Stuart, 2004). Matched cases were selected based on 1:1 nearest neighbor matching. A caliper was used that required the standardized mean difference between the matched treatment and comparison cases to be at or below the specified value of 0.10 standard deviations on all the child and family covariates that were included in the matching process (see Table 1). Cases were randomly ordered within groups prior to matching, and cases that did not have good matches were discarded. The 1:1 matching procedure resulted in a sample of 532 children in the treatment (n = 266) and comparison groups (n = 266).

Regression analyses

A series of multi-level regression models was used to examine the effects of NC Pre-K participation (i.e., the treatment) on child outcomes at the end of kindergarten as well as moderation of the NC Pre-K effects by measures of elementary school-wide quality. Multi-level regression models were estimated using the MIXED procedure in SAS (Version 9.4) with restricted maximum likelihood (ReML) estimation. Separate models were estimated for language, literacy, mathematics, and short-term working memory outcomes. We accounted for the nesting of children within pre-K classrooms, because this is the context in which the NC Pre-K treatment was administered, and because the focus of our analyses concerned NC Pre-K effects and moderation of these effects.

In these two-level models, children (Level-1) were nested within pre-K classrooms (Level-2). To account for the partially nested structure of children nested in pre-K classrooms,[3] the variance of the Level-1 residuals was allowed to differ across the NC Pre-K and comparison groups (see Bauer, Sterba, & Hallfors, 2008 for details). The following generalized equation was used:

$$\begin{aligned} Y_{ij} = {} & \gamma_{00} + \gamma_{10}\,\text{treatment}_j + \gamma_{20}\,\text{academic proficiency}_{ij} + \gamma_{30}\,\text{academic growth}_{ij} \\ & + \gamma_{40}\,\text{school poverty}_{ij} + \gamma_{50}\,\text{school size}_{ij} + \gamma_{60}\,\text{treatment} \\ & \times \text{academic proficiency}_{ij} + \gamma_{70}\,\text{treatment} \times \text{academic growth}_{ij} \\ & + \gamma_{80}\,\text{treatment} \times \text{school poverty}_{ij} + \gamma_{90}\,\text{treatment} \times \text{school size}_{ij} \\ & + \gamma_{100}\,\text{covariate}_{ij} + \ldots + \gamma_{190}\,\text{covariate}_{ij} + u_j + r_{ij}. \end{aligned}$$

In this generalized equation, Y_{ij} is an outcome score for child i in the pre-K classroom j; γ_{00} is the intercept representing the mean of child outcome scores; γ_{10} is the main effect of *treatment* (0 = Comparison group, 1 = NC Pre-K group) at the average level of school quality; γ_{20} is the main effect of *academic proficiency*; γ_{30} is the main effect of *academic growth*; γ_{40} is the main effect of *school poverty*; γ_{50} is the main effect of *school size*; γ_{60} is the interaction between *treatment* and *academic proficiency*; γ_{70} is the interaction between *treatment* and *academic growth*; γ_{80} is the interaction between *treatment* and *school poverty*; γ_{90} is the interaction between *treatment* and *school size*; γ_{100} through γ_{190} are the main effects of the child and family background covariates (i.e., gender, Latino ethnicity, race, chronic health condition, English spoken at home, parent in military, primary caregiver education, and family size); u_j is the Level-2 random effect that permits the mean of the outcome scores to vary across pre-K classrooms; r_{ij} is the Level-1 residual that is allowed to have different variances across the NC Pre-K and comparison groups. We employed a method termed *doubly robust* by including in the regression analyses many of the child and family background covariates that were also included in the PSM analyses, in order to remove additional sources of bias in estimating the NC Pre-K effects. Three modifications were made to the set of covariates that were included in the PSM analyses. Given the small number of children included in some of the racial categories, the separate race variables were replaced with a White/non-White variable. Because of high inter-correlations and concerns about multicollinearity, three language variables were dropped (child speaks English at home, English is child's first language, and whether the survey was completed in English), as they were highly correlated with the frequency the primary caregiver speaks English at home (r = 0.84–0.85). In addition, of the three variables related to household composition, we retained family size but excluded the number of adults and number of

[3] We chose to account for the partial nesting structure of the data because a greater percent of children in the treatment group were enrolled in the same pre-K classroom as another treatment group child (67%) compared to the percent of children in the comparison group who were enrolled in the same pre-K classroom as another comparison group child (9%). Children in the comparison group who were never enrolled in a formal pre-K classroom were categorized with their own unique value for the pre-K classroom membership variable.

Table 1
Treatment and comparison group means before and after propensity score matching.

	Before Propensity Score Matching			After Propensity Score Matching		
Covariates	NC Pre-K (n = 432)	Comparison (n = 391)	SMD	NC Pre-K (n = 266)	Comparison (n = 266)	SMD
Child Characteristics						
Gender	0.51	0.46	0.12	0.47	0.48	−0.02
Latino ethnicity	0.33	0.29	0.09	0.27	0.30	−0.08
Race						
African American	0.32	0.30	0.05	0.35	0.32	0.07
American Indian/Alaska Native	0.05	0.04	0.04	0.04	0.05	−0.05
Asian	0.01	0.02	−0.13	0.01	0.02	−0.04
Chronic health condition	0.11	0.13	−0.05	0.13	0.13	0.01
Child's first language is English	0.70	0.77	−0.15	0.78	0.76	0.03
Child speaks English at home	0.69	0.75	−0.11	0.77	0.75	0.04
Family Characteristics						
English spoken at home	3.33	3.46	−0.11	3.50	3.47	0.03
Military parent	0.16	0.11	0.13	0.11	0.12	−0.03
Primary Caregiver Education	2.96	2.91	0.05	2.98	2.99	0.00
Family size	4.41	4.50	−0.07	4.32	4.45	−0.10
Number of adults in the home	1.94	1.94	0.00	1.92	1.96	−0.05
Number of children in the home	2.53	2.68	−0.14	2.49	2.61	−0.10
Family income	1.69	1.71	−0.02	1.68	1.72	−0.03
Survey completed in English	0.79	0.85	−0.16	0.86	0.85	0.05

Note.\ SMD = Standardized mean difference.

Table 2
Descriptive statistics on school quality measures and school covariates.

Measures	State		Sample		t-test
	n = 1558		n = 33		
	M	SD	M	SD	
Academic proficiency (%)	0.60	0.16	0.55	0.16	t (1589) = 1.78, p = .08
Academic growth	0.00	1.84	0.22	3.13	t (1589) = 0.67, p = .50
School poverty (%)	0.66	0.29	0.75	0.20	t (1589) = 1.77, p = .07
School size	473	358	494	213	t (1589) = 0.34, p = .74

Note. Information about state-level descriptive statistics was calculated by the authors using data obtained from the North Carolina Department of Public Instruction, Accountability and Services Division. State information was based on all schools that served an elementary school grade (i.e., kindergarten, first, second, third, fourth, and/or fifth).

children as the family size variable is the sum of those two excluded variables. All the continuous independent and dependent variables were standardized to have a mean of 0 and standard deviation of 1 in order to report standardized regression coefficients in Table 4, which can be interpreted as effect sizes (Hedges, 2008), while the dichotomous predictor variables were mean centered, but not standardized (i.e., treatment, gender, Latino ethnicity, race, chronic health condition, parent in military, and teaching license). Standardizing and centering also provides comparable main effect coefficients to models without interactions.

Missing data

From the matched sample of 532 children (n = 266 treatment and n = 266 comparison), child outcome data could not be gathered for 20 children who left the school before child assessments were administered at the end of kindergarten, resulting in an attrition rate of 3.7%, which is considered negligible in educational research (Peugh & Enders, 2004). Children with missing outcome data were excluded from the analyses using listwise deletion, which resulted in an analytic sample of 512 children across the treatment (n = 255) and comparison (n = 257) groups for the language and short-term working memory outcome measures. One or two children were also missing data for the literacy and mathematics measures, respectively, because they were unable to complete some of the subtests.

Analyses

Descriptive statistics

The current study included an analytic sample of 512 children (n = 255 treatment, n = 257 comparison) who attended 33 elementary schools (see Table 2). We conducted t-tests to compare the school characteristics between the sample of 33 elementary schools in our dataset and the state-wide average. These findings indicated that the characteristics of the elementary schools experienced by children in our sample were not reliably different from the state-wide average. In terms of school-wide academic proficiency, in the average elementary school in our sample, 55% of students scored in the proficient range on the EOG reading and math tests in third grade, which was not reliably different from the state-wide average of 60% (t [1589] = 1.78, p = .08). The average school-wide academic growth score in our sample was 0.22 (SD = 3.13), which was not reliably different from the state-wide average of 0.00 (SD = 1.84) (t [1589] = 0.67, p = .50). The percent of economically disadvantaged students in the school (i.e., school poverty) was 75% on average in our sample, which was not reliably different from the state-wide average of 66% (t [1589] = 1.77, p = .07). On average, the schools in our sample had 494 (SD = 213) students enrolled, which was not reliably different from the state-wide average of 473 (SD = 358) (t [1589] = 0.34, p = .74).

Table 3
Correlations between treatment, school quality measures, school covariates, and child and family covariates (n = 512).

		(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)	(k)	(l)	(m)	(n)
a.	Treatment (1 = NC Pre-K)	1.00													
b.	Academic proficiency	0.00	1.00												
c.	Academic growth	0.05	0.19	1.00											
d.	School poverty	0.05	−0.30	0.33	1.00										
e.	School size	−0.01	−0.17	0.03	−0.24	1.00									
f.	Gender (1 = Male)	0.01	−0.06	−0.03	0.02	−0.04	1.00								
g.	Latino ethnicity (1 = Yes)	−0.04	−0.08	0.15	0.16	0.15	0.02	1.00							
h.	Race (1 = White)	0.00	0.18	−0.11	−0.21	−0.15	0.05	−0.22	1.00						
i.	Chronic health condition (1 = Yes)	0.00	−0.03	−0.02	0.03	−0.02	0.03	−0.08	−0.07	1.00					
j.	English spoken at home	0.01	0.09	−0.14	−0.11	−0.14	0.00	−0.74	0.29	0.11	1.00				
k.	Parent in military (1 = Yes)	−0.02	−0.06	−0.17	−0.06	0.09	−0.03	0.09	−0.06	−0.10	−0.12	1.00			
l.	Primary caregiver education	0.00	0.08	−0.12	−0.21	−0.06	−0.04	−0.43	0.17	0.02	0.46	−0.05	1.00		
m.	Family size	−0.06	0.04	−0.04	0.00	−0.01	−0.05	0.10	−0.04	−0.07	−0.14	0.08	−0.11	1.00	
n.	Family income	−0.03	0.15	−0.22	−0.32	0.01	0.02	−0.13	0.20	−0.08	0.11	0.11	0.26	0.23	1.00

Note. Values greater than or equal to 0.09 are statistically significant below the $p < .05$ level.

Correlations

We examined bivariate correlations between all the independent variables in our analyses (see Table 3). Treatment (i.e., participation in NC Pre-K) was not highly correlated with the child and family covariates ($r = 0.00$ to $|0.06|$), nor the school quality measures ($r = 0.00$ and 0.05) nor school covariates ($r = -0.01$ and 0.05). Bivariate correlations between the child and family covariates were generally small to moderate ($r = 0.00$ to $|0.29|$), with high correlations in a few cases among primary caregiver education, English spoken at home, and Latino ethnicity ($r = 0.43$ to $|0.74|$). The bivariate correlation between school-wide academic proficiency and academic growth was moderate ($r = 0.19$), and bivariate correlations between the school quality measures, school covariates, and child/family covariates ranged from small to moderate ($r = 0.00$ to $|0.33|$).

Regression results

We examined the effects of NC Pre-K participation (i.e., the treatment) on each child outcome measure, which was estimated at the average level of school quality. We simultaneously examined moderation of the NC Pre-K effects by the measures of elementary school-wide quality. Results for the regression analyses are displayed in Table 4.

Language

The effect of NC Pre-K participation on children's language skills was not statistically significant at the average level of school quality ($\beta = 0.08$, $p = .28$), but was moderated by school-wide academic proficiency. Moderation analyses indicated a 0.16 increase in the effect of NC Pre-K participation with every 1 standard deviation (*SD*) unit increase in academic proficiency ($\beta = 0.16$, $p = .049$; see Fig. 1 in which we plotted the effect at varying levels of academic proficiency between −1 and + 1 *SD* of the average). Probing the interaction revealed that this effect was marginally significant in schools rated at +0.5 *SD* of the average on academic proficiency ($\beta = 0.16$, $p = .06$) and statistically significant in schools rated at +1 *SD* of the average ($\beta = 0.24$, $p = .03$). No evidence of moderation was found for school-wide academic growth.

Literacy

The effect of NC Pre-K participation on children's literacy skills was not statistically significant at the average level of school quality ($\beta = -0.02$, $p = .81$) and no evidence of moderation was found for school-wide academic proficiency or academic growth.

Mathematics

The effect of NC Pre-K participation on children's mathematics skills was not statistically significant at the average level of school quality ($\beta = 0.14$, $p = .12$) and no evidence of moderation was found for school-

Table 4
Regression results.

Parameters	Language		Literacy		Mathematics		Short-Term Working Memory	
	n = 512		*n* = 511		*n* = 510		*n* = 512	
	β	(*SE*)	β	(*SE*)	β	(*SE*)	β	(*SE*)
Treatment (1 = NC Pre-K)	0.08	(0.07)	−0.02	(0.09)	0.14	(0.09)	0.16	(0.08)
School-wide academic proficiency	−0.05	(0.04)	0.01	(0.05)	0.06	(0.05)	0.00	(0.05)
School-wide academic growth	0.06	(0.04)	−0.03	(0.05)	−0.04	(0.05)	−0.05	(0.05)
School poverty	−0.11***	(0.05)	0.01	(0.06)	0.06	(0.06)	0.04	(0.06)
School size	0.01	(0.04)	0.04	(0.05)	0.11*	(0.05)	0.05	(0.05)
Treatment × academic proficiency	0.16*	(0.08)	0.11	(0.11)	0.10	(0.11)	0.08	(0.10)
Treatment × academic growth	−0.01	(0.08)	−0.09	(0.11)	−0.04	(0.11)	0.20*	(0.10)
Treatment × school poverty	0.06	(0.09)	0.07	(0.11)	−0.03	(0.12)	−0.02	(0.11)
Treatment × school size	0.03	(0.08)	0.06	(0.10)	0.04	(0.10)	0.09	(0.09)
Variance Components	*Estimate*		*Estimate*		*Estimate*		*Estimate*	
Level-1: Comparison group	0.56***		0.79***		0.72***		1.01***	
Level-1: NC Pre-K group	0.44***		0.83***		0.80***		0.81***	
Level-2	0.27		0.52		0.72*		0.00	

Note. Parameters γ_{100} through γ_{190} were included in all models, but their estimates were not displayed in Table 4 and can be made available from the corresponding author upon request.
* $p < .05$.
** $p < .01$.
*** $p < .001$.

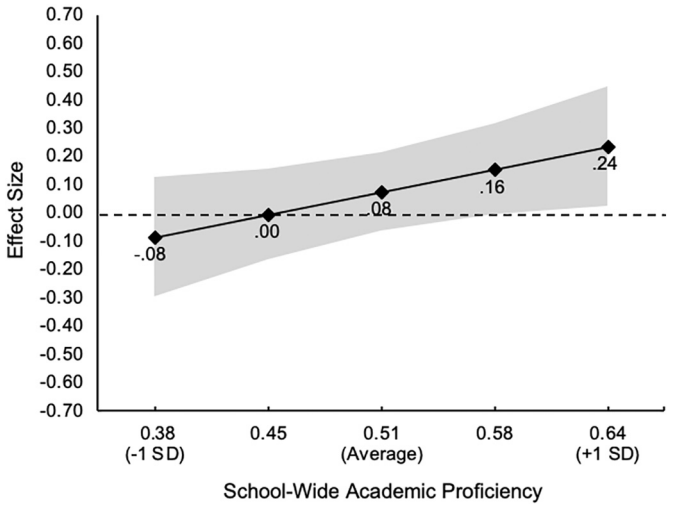


Fig. 1. The effect of NC Pre-K participation on children's language skills moderated by school-wide academic proficiency.
Note. Effect sizes are reported as standardized regression coefficients (β). The shaded region represents the 95% confidence interval. *SD* = standard deviation.

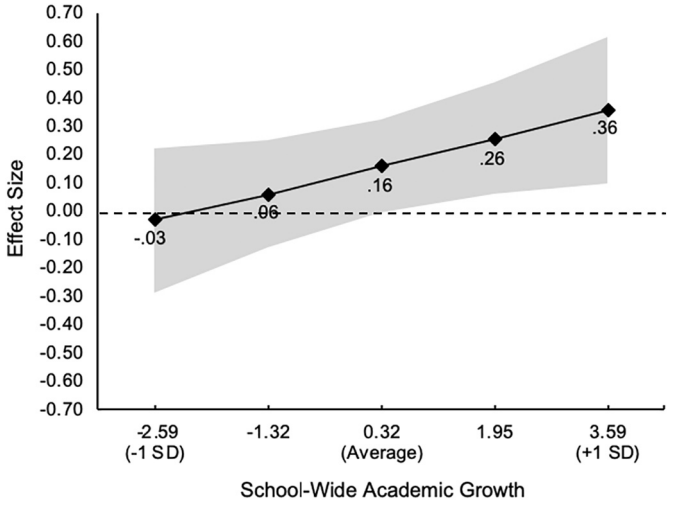


Fig. 2. The effect of NC Pre-K participation on children's short-term working memory skills moderated by school-wide academic growth.
Note. Effect sizes are reported as standardized regression coefficients (β). The shaded region represents the 95% confidence interval. *SD* = standard deviation.

wide academic proficiency or academic growth.

Short-term working memory

The effect of NC Pre-K participation on children's short-term working memory skills was marginally significant at the average level of school quality (β = 0.16, p = .06) and was moderated by school-wide academic growth. Moderation analyses indicated a 0.20 increase in the effect with every 1 *SD* unit increase in academic growth (β = 0.20, p = .047; see Fig. 2 in which we plotted the effect at varying levels of academic growth between −1 and + 1 *SD* of the average). Probing the interaction revealed that this effect was statistically significant in schools rated at +0.5 *SD* of the average level of academic growth (β = 0.26, p = .01) and + 1 *SD* of the average (β = 0.36, p = .006). No evidence of moderation was found for school-wide academic proficiency.

Sensitivity analyses

In a series of sensitivity analyses, we examined if both of the statistically significant moderating effects found in our primary analyses remained significant in models excluding other school-level measures, but retaining all child-level covariates, because issues associated with multicollinearity could result as a consequence of simultaneously including these moderately correlated measures in the same model. Our sensitivity analyses revealed that the moderating effect of academic proficiency on children's language skills was slightly diminished in magnitude from the full model (β = 0.16, *SE* = 0.08, p = .049) to the reduced model (β = 0.14, *SE* = 0.07, p = .047), but remained statistically significant in the reduced model. We also found that the moderating effect of academic growth on children's short-term working memory skills was slightly enhanced in magnitude from the full model (β = 0.20, *SE* = 0.10, p = .047) to the reduced model (β = 0.21, *SE* = 0.09, p = .014), and remained statistically significant in the reduced model. Additionally, we considered a series of sensitivity analyses in which quadratic terms for school-wide academic proficiency and academic growth were included as moderators of the treatment effects and we found no evidence that the moderating effects of school quality operated in a non-linear manner.

Discussion

The favorable effects of pre-K programming may diminish over time unless children continue to experience high-quality educational environments in elementary school (Bailey et al., 2017; Phillips et al., 2017). However, few studies have examined the contributions of elementary school quality in relation to the effects of pre-K programming, and those studies provide mixed results (Bailey et al., 2020). In the current study, we examined the extent to which NC Pre-K program participation affected child outcomes at the end of kindergarten as well as the extent to which these effects were moderated by measures of school-wide quality.

The current study did not find reliable effects of NC Pre-K participation on children's language, literacy, mathematics, and working memory skills at the average level of school quality. However, measures of school-wide academic proficiency and school-wide academic growth moderated the effects of NC Pre-K participation on children's language and working memory skills, respectively. The favorable effect of NC Pre-K on children's language skills was only found in schools rated above-average on academic proficiency (measured by the percent of third-grade students in the school who scored in the proficient range on standardized assessments of reading and mathematics). Similarly, the favorable effect of NC Pre-K on children's working memory skills was only found in schools rated above-average on academic growth (measured by the school-wide rate of student growth on standardized assessments across grades for all students in third grade through the school-max grade). When considered together, these findings suggest that school-wide quality played a role in enhancing the effects of NC Pre-K participation on children's language and working memory skills at the end of kindergarten. Moreover, these findings provided evidence of a dynamic complementarity between NC Pre-K participation and elementary school quality.

School-wide academic proficiency and growth are commonly used as policy metrics of school quality. However, developmental theory related to peer learning and peer environments can explain how these measures may enhance pre-K effects. First, the academic peer environment of a school may provide greater opportunities for co-learning with peers. School environments with high-proficient and/or high-growth peers may cultivate positive academic behavior, shared social norms valuing academic engagement and success, or the absence of chaotic learning environments (Cialdini & Trost, 1998; Eccles & Roeser, 2010; Hanushek et al., 2003; Lavy et al., 2012; Lynch et al., 2013). Second, the academic peer environment of a school may impact the instruction provided by teachers and, in turn, the academic achievement of students (Henry & Rickman, 2007). Unfortunately, our study was not able to test the specific mechanisms through which the peer environment may enhance the

effects of NC Pre-K. However, it is reasonable to expect that children who graduate from a high-quality pre-K program will be more likely to benefit from an elementary school context where the student body demonstrates higher levels of academic proficiency and/or growth.

Given that the pattern of moderation findings was varied (i.e., school-wide academic proficiency moderated the NC Pre-K effect on child language, while school-wide academic growth moderated the NC Pre-K effect on child working memory), it is necessary to speculate about the ways in which these different dimensions of school quality may contribute to enhancing NC Pre-K effects on different outcomes. In relation to language, an academic peer environment characterized by high-proficient peers may also be an environment in which more complex language is being used to teach and learn academic content. Therefore, former NC Pre-K participants may rapidly acquire more advanced language in this environment—enhancing the NC Pre-K effect on child language skills. In relation to working memory, an academic peer environment characterized by high-growth peers may be one in which the basic skills and dispositions of learning are emphasized, independent of any specific content area. Moreover, our assessments of working memory were relatively content free, unlike our assessment of language, which may have been closely linked to the curricular content and concepts being taught in the classroom. Therefore, former NC Pre-K participants may be able to more rapidly develop foundational learning skills, such as working memory skills, in academic peer environments characterized by high-growth peers.

Prior research has documented positive short-term effects of NC Pre-K on children's school readiness skills (Peisner-Feinberg et al., 2019; Peisner-Feinberg & Schaaf, 2011) as well as positive longer-term effects on children's academic progress during elementary school and beyond (Bai et al., 2020; Dodge et al., 2016; Ladd et al., 2014; Muschkin et al., 2015; Peisner-Feinberg & Schaaf, 2010). In contrast to prior studies, we were surprised to find no reliable effects of NC Pre-K participation on children's literacy or mathematics skills in elementary school. However, we did find positive effects of NC Pre-K participation on children's end-of-kindergarten language and working memory skills in the context of higher-quality elementary schools. It was encouraging to find positive effects on these "unconstrained" skills, because children may possess unlimited potential for growth in these skills during the early elementary grades. Alternatively, the measures of literacy and mathematics used in this study capture the development of more "constrained" skills that all children are expected to develop during the early elementary grades (McCormick, Hsueh, Weiland, & Bangser, 2017). It is possible that the effect of NC Pre-K programming on children's unconstrained skills at the end of kindergarten would translate into longer-term effects in later elementary grades on academic skills such as reading and mathematics. This may be the case because skills related to working memory, for example, entail domain-general cognitive processes that facilitate problem solving in academic domains of learning. Prior research provides support for this hypothesis, as lead-lag associations have been widely documented between children's working memory skills and subsequent academic skills during the early elementary school grades (e.g., Willoughby, Wylie, & Little, 2019).

The current study examined the combined effects of NC Pre-K and subsequent elementary school quality. While our findings were mixed, findings from prior research into factors that enhance the effects of ECE programming during later grades are also notably mixed (Bailey et al., 2020). However, our findings are consistent with a previous study that documented an enhanced effect of the Chicago School Readiness Project on language/literacy skills at the end of kindergarten for children who subsequently entered highly proficient elementary schools (i.e., 0.5 standard deviations above the average; Zhai et al., 2012). While many previous studies have examined measures of school-wide academic proficiency in relation to the effects of ECE programming, our study was one of a handful to examine school-wide academic growth. Our findings related to school-wide academic growth and working memory were similar to recent studies that linked school-wide academic growth to the enhanced effects of publicly funded pre-K programming on third grade reading and/or mathematics outcomes (Pearman et al., 2020; Unterman & Weiland, 2020).

Implications

Policymakers and education stakeholders recognize the importance of aligning educational systems and inputs across the preschool through third grade continuum to support children's optimal learning and development during this period. For example, the National P-3 Center has developed a framework "to improve the quality and coherence of children's learning opportunities" across preschool through third grade (Kauerz & Coffman, 2013, p. 1). Although there are only a handful of states beginning to implement preschool–third grade alignment policies and programs at scale (www.nationalp-3center.org/map), NC has several state-level initiatives that have been undertaken to inform public policymaking related to this issue. For example, the B-3 Interagency Council was established by the NC House of Representatives and charged with planning "a coordinated system of early care, education, and child development services… in an effort to optimize learning gains realized in the prekindergarten years" (Section 7.23I. (a); Current Operations and Appropriations Act of 2017, 2017; p. 59), and the Early Childhood Advisory Council was established by NC Governor Roy Cooper and charged to coordinate with the B-3 Interagency Council on this plan (NC Executive Order No. 40, 2018). Additionally, many local community organizations across NC endorsed the NC Pathways to Grade-Level Reading initiative, which proposed to expand access to NC Pre-K and improve elementary school quality as strategies to improve high-quality birth-through-age-eight learning environments (NC Early Learning Foundation, 2017).

Results from the current study suggest that NC Pre-K programming and its benefits should be viewed as part of a continuum of high-quality educational experience to be built upon sequentially across the transition from preschool to elementary school. We also identify school-wide academic proficiency and growth as two types of quality inputs in elementary school that may productively build on the quality inputs children receive through participation in NC Pre-K. These findings also offer implications for the alignment of high-quality educational experiences across pre-K and elementary education more broadly. First, it is important to acknowledge that efforts to foster alignment across education sectors may be difficult to implement in many states, including NC, where pre-K programming operates independently from the public-school system. The current study identified two measures of elementary school quality (school-wide academic proficiency and academic growth) that public-school administrators may focus on to enhance the effects of the high-quality NC Pre-K program—independent of coordination with the NC Pre-K program. Moreover, education stakeholders and policymakers in NC and nationally are already focused on improving these metrics of school quality, because they are used for federal accountability and school reform initiatives. Second, measures of school-wide quality should be devised to assess the academic proficiency and academic growth of students in kindergarten through second grade. Given that the measures of school quality considered in this study were based on the test scores of students beginning in third grade, such metrics may not perfectly represent the quality of the school environment experienced by children in earlier elementary grades. Moreover, measuring school quality based on kindergarten through second grade academic proficiency and growth may promote better alignment of high-quality educational environments across pre-K and early elementary school. Finally, it will be important for future research to consider measures of both classroom-level and school-level quality in relation to the effects of pre-K programming, because classroom-level quality has also been found to enhance pre-K effects during kindergarten (Carr, Mokrova, Vernon-Feagans, & Burchinal, 2019).

Limitations

The results of our study should be viewed in light of certain limitations. First, propensity score matching cannot adjust for unmeasured confounding to the same extent as a randomized control trial. Therefore, our analyses may have been biased by unobserved differences between the NC Pre-K and comparison groups. Second, because this study only examined child outcomes at the end of kindergarten, we could not consider the initial effects of NC Pre-K participation at the end of pre-K. Therefore, our analyses could not distinguish between (a) effects that were evident at the end of pre-K and enhanced in high-quality school contexts during kindergarten, and (b) effects that were not evident at the end of pre-K but emerged in high-quality contexts. Nonetheless, this does not discount the differential effect of NC Pre-K based on school-wide academic proficiency and growth. We also note that there is a long history of statewide evaluations documenting high-quality NC Pre-K programming and robust findings to indicate that NC Pre-K participation has favorable short-term and long-term effects, as reviewed in earlier sections. Third, unlike some other studies of pre-K program effects, we chose to exclude children with IEPs and identified disabilities from this sample due to concerns that the demands of the extensive assessment battery and assessment protocol would be inappropriate for these children. This decision may have downwardly biased our estimates of NC Pre-K effects given that prior research has documented positive pre-K program effects on academic outcomes for children with disabilities (Weiland, 2016). Finally, our study did not consider other dimensions of school quality, such as principal leadership or school climate, which may be related to the measures of school quality and child outcomes considered in this study. However, we chose to focus on school-wide academic proficiency and academic growth, given their widespread use in school accountability and reform efforts.

Conclusion

The current study contributes new research to a growing body of evidence on factors that enhance the effects of public pre-K programming during elementary school. We focused on the NC Pre-K program—an established, high-quality pre-K program. The results of this study suggest that school-wide quality in elementary school played a role in enhancing the effects of NC Pre-K through the end of kindergarten—one year after children had finished participating in the program. Therefore, we advocate for a view of pre-K programming and its potential benefits as part of a continuum of high-quality educational experience to be built on sequentially across the transition to elementary school. Moreover, we advocate for efforts to ensure continuity in high-quality learning environments throughout this period. Our study identifies two common metrics of school quality that may make a difference in enhancing the benefits of NC Pre-K programming during elementary school: school-wide academic proficiency and academic growth. While these metrics are valuable starting points both for policy and research, we suggest that future research also focus in on measures specifically designed to assess academic proficiency and academic growth among K–2 students, along with measures of classroom-level quality in the early elementary grades.

Funding

This research study was supported by a Pre-doctoral State Policy Fellowship in Early Learning provided by the Society for Research in Child Development to the first author (R.C.). The original study was supported by a contract from the North Carolina Division of Child Development and Early Education (DCDEE), North Carolina Department of Health and Human Services (DHHS) to the second author (E.P-F.). The authors wish to acknowledge that some ideas presented in this manuscript were formulated during a convening of the Consortium on Early Childhood Intervention Impact (CECII), which was made possible by grant funding from the National Institute of Child Health and Human Development (1R01HD095930-01A1).

Declaration of Competing Interest

None.

Acknowledgements

We express our gratitude to all those who participated in and assisted with this study, including the children, families, teachers, principals, administrators, and other staff in the kindergarten programs and the staff of DCDEE. Thank you to Helen Ladd for her helpful feedback on this manuscript.

References

Ansari, A., & Pianta, R. C. (2018). The role of elementary school quality in the persistence of preschool effects. *Children and Youth Services Review, 86*(1), 120–127. https://doi.org/10.1016/j.childyouth.2018.01.025

Bai, Y., Ladd, H. F., Muschkin, C. G., & Dodge, K. A. (2020). Long-term effects of early childhood programs through eighth grade: Do the effects fade out or grow? *Children and Youth Services Review, 112*(1). https://doi.org/10.1016/j.childyouth.2020.104890

Bailey, D., Duncan, G. J., Odgers, C. L., & Yu, W. (2017). Persistence and fadeout in the impacts of child and adolescent interventions. *Journal of Research on Educational Effectiveness, 10*(1), 7–39. https://doi.org/10.1080/19345747.2016.1232459

Bailey, D. H., Jenkins, J. M., & Alvarez-Vargas, D. (2020). Complementarities between early educational intervention and later educational quality? A systematic review of the sustaining environments hypothesis. *Developmental Review, 56*(1). https://doi.org/10.1016/j.dr.2020.100910

Barnett, S. W., Friedman-Krauss, A. H., Gomez, R., Horowitz, M., Weisenfeld, G. G., Clarke Brown, K., & Squires, J. H. (2015). *The state of preschool 2015: State preschool yearbook*. National Institute of Early Education Research. www.nieer.org.

Barnett, W. S., Jung, K., Friedman-Krauss, A., Frede, E. C., Nores, M., Hustedt, J. T., … Daniel-Echols, M. (2018). State prekindergarten effects on early learning at kindergarten entry: An analysis of eight state programs. *AERA Open, 4*(2), 1–16. https://doi.org/10.1177/2332858418766291

Bassok, D., Gibbs, C. R., & Latham, S. (2018). Preschool and children's outcomes in elementary school: Have patterns changed nationwide between 1998 and 2010? *Child Development, 90*(6), 1–23. https://doi.org/10.1111/cdev.13067

Battistich, V., Solomon, D., Kim, D.-I., Watson, M., & Schaps, E. (1995). Schools as communities, poverty levels of student populations, and students' attitudes, motives, and performance: A multilevel analysis. *American Educational Research Journal, 32*(3), 627–658. https://doi.org/10.3102/00028312032003627

Bauer, D. J., Sterba, S. K., & Hallfors, D. D. (2008). Evaluating group-based interventions when control participants are ungrouped. *Multivariate Behavioral Research, 43*(2), 210–236. https://doi.org/10.1080/00273170802034810

Bierman, K. L., Nix, R. L., Heinrichs, B. S., Domitrovich, C. E., Gest, S. D., Welsh, J. A., & Gill, S. (2014). Effects of head start REDI on children's outcomes 1 year later in different kindergarten contexts. *Child Development, 85*(1), 140–159. https://doi.org/10.1111/cdev.12117

Carr, R. C., Mokrova, I. L., Vernon-Feagans, L., & Burchinal, M. R. (2019). Cumulative classroom quality during pre-kindergarten and kindergarten and children's language, literacy, and mathematics skills. *Early Childhood Research Quarterly, 47*(2), 218–228. https://doi.org/10.1016/j.ecresq.2018.12.010

Cialdini, R. B., & Trost, M. R. (1998). Social influence: Social norms, conformity and compliance. In D. T. Gilbert, S. T. Fiske, et al. (Eds.) (4th ed.,, *Vol. 2. The handbook of social psychology* (pp. 151–192). McGraw-Hill.

Cohen-Vogel, L., Sadler, J., Little, M. H., Merrill, B., & Curran, F. C. (2020). The adoption of public pre-kindergarten among the American states: An event history analysis. *Educational Policy*, 1–33. https://doi.org/10.1177/0895904820961002. Advance online publication.

Cunha, F., & Heckman, J. (2007). The technology of skill formation. *The American Economic Review, 97*(2), 31–47. https://doi.org/10.1257/aer.97.2.31

Curenton, S. M., Dong, N., & Shen, X. (2015). Does aggregate school-wide achievement mediate fifth grade outcomes for former early childhood education participants? *Developmental Psychology, 51*(7), 921–934. https://doi.org/10.1037/a0039295

Dodge, K. A., Bai, Y., Ladd, H. F., & Muschkin, C. G. (2016). Impact of North Carolina's early childhood programs and policies on educational outcomes in elementary school. *Child Development, 88*(3), 996–1014. https://doi.org/10.1111/cdev.12645

Eccles, J. S., & Roeser, R. W. (2010). An ecological view of schools and development. In J. S. Eccles, & J. L. Meece (Eds.), *Handbook of research on schools, schooling, and human development* (pp. 6–21). Routledge.

Friedman-Krauss, A. H., Barnett, S. W., Garver, K. A., Hodges, K. S., Weisenfeld, G. G., & DiCrecchio, N. (2019). The state of preschool 2018: State preschool yearbook. National Institute of Early Education Research www.nieer.org.

Fuhs, M. E., & Day, J. D. (2011). Verbal ability and executive functioning development in preschoolers at head start. *Developmental Psychology, 47*(2), 404–416.

Gathercole, S. E., & Pickering, S. J. (2000). Working memory deficits in children with low achievements in the national curriculum at 7 years of age. *British Journal of*

Educational Psychology, 70(1), 177–194. https://doi.org/10.1348/000709900158047
Gershenson, S., & Langbein, L. (2015). The effect of primary school size on academic achievement. *Educational Evaluation and Policy Analysis, 37*(1), 135–155. https://doi.org/10.3102/0162373715576075
Gormley, W. T., Phillips, D., & Anderson, S. (2018). The effects of Tulsa's Pre-K program on middle school student performance. *Journal of Policy Analysis and Management, 37*(1), 63–87. https://doi.org/10.1002/pam.22023
Hanushek, E. A., Kain, J. F., Markman, J. M., & Rivkin, S. G. (2003). Does peer ability affect student achievement? *Journal of Applied Econometrics, 18*(5), 527–544. https://doi.org/10.1002/jae.741
Harms, T., Clifford, R. M., & Cryer, D. (1998). *Early childhood environment rating scale–revised*. Teachers College Press.
Hedges, L. V. (2008). What are effect sizes and why do we need them? *Child Development Perspectives, 2*(3), 167–171. https://doi.org/10.1111/j.1750-8606.2008.00060.x
Henry, G. T., & Rickman, D. K. (2007). Do peers influence children's skill development in preschool? *Economics of Education Review, 26*(1), 100–112. https://doi.org/10.1016/j.econedurev.2005.09.006
Hill, C. J., Gormley, W. T., & Adelstein, S. (2015). Do the short-term effects of a high-quality preschool program persist? *Early Childhood Research Quarterly, 32*(3), 60–79. https://doi.org/10.1016/j.ecresq.2014.12.005
Ho, D., Imai, K., King, G., & Stuart, E. (2004). Matchit: Matching as nonparametric preprocessing for parametric causal inference. *Journal of Statistical Software, 42*(8), 1–28. https://doi.org/10.18637/jss.v042.i08
Jenkins, J. M., Watts, T. W., Magnuson, K., Gershoff, E., Clements, D., Sarama, J., & Duncan, G. J. (2018). Do high quality kindergarten and first grade classrooms mitigate preschool fadeout? *Journal of Research on Educational Effectiveness, 11*(3), 339–374. https://doi.org/10.1080/19345747.2018.1441347
Kauerz, K., & Coffman, J. (2013). *Framework for planning, implementing, and evaluating prek-3rd grade approaches*. National P-3 Center. www.nationalp-3center.org.
Kuziemko, I. (2006). Using shocks to school enrollment to estimate the effect of school size on student achievement. *Economics of Education Review, 25*(1), 63–75. https://doi.org/10.1016/j.econedurev.2004.10.003
Ladd, H. F., Muschkin, C. G., & Dodge, K. A. (2014). From birth to school: Early childhood initiatives and third-grade outcomes in North Carolina. *Journal of Policy Analysis and Management, 33*(1), 162–187. https://doi.org/10.1002/pam.21734
Lavy, V., Paserman, M. D., & Schlosser, A. (2012). Inside the black box of ability peer effects: Evidence from variation in the proportion of low achievers in the classroom. *The Economic Journal, 122*(559), 208–237. https://doi.org/10.1111/j.1468-0297.2011.02463.x
Li, W., Leak, J., Duncan, G. J., Magnuson, K., Schindler, H., & Yoshikawa, H. (2016). *Is timing everything? How early childhood education program impacts vary by starting age, program duration and time since the end of the program* (Working paper).
Lipsey, M. W., Farran, D. C., & Durkin, K. (2018). Effects of the Tennessee prekindergarten program on children's achievement and behavior through third grade. *Early Childhood Research Quarterly, 45*(1), 155–176. https://doi.org/10.1016/j.ecresq.2018.03.005
Lipsey, M. W., Nesbitt, K. T., Farran, D. C., Dong, N., Fuhs, M. W., & Wilson, S. J. (2017). Learning-related cognitive self-regulation measures for prekindergarten children: A comparative evaluation of the educational relevance of selected measures. *Journal of Educational Psychology, 109*(8), 1084–1102. https://doi.org/10.1037/edu0000203
Lynch, A. D., Lerner, R. M., & Leventhal, T. (2013). Adolescent academic achievement and school engagement: An examination of the role of school-wide peer culture. *Journal of Youth and Adolescence, 42*(1), 6–19. https://doi.org/10.1007/s10964-012-9833-0
McCormick, M., Hsueh, J., Weiland, C., & Bangser, M. (2017). The challenge of sustaining preschool impacts: Introducing ExCEL P-3, a study from the expanding Children's Early Learning Network. MDRC www.mdrc.org.
McCoy, D. C., Roy, A. L., & Sirkman, G. M. (2013). Neighborhood crime and school climate as predictors of elementary school academic quality: A cross-lagged panel analysis. *American Journal of Community Psychology, 52*(1–2), 128–140. https://doi.org/10.1007/s10464-013-9583-5
McGrew, K. S., & Woodcock, R. W. (2001). *Woodcock-Johnson III technical manual*. The Riverside Publishing Company.
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*(4), 478–500. https://doi.org/10.3102/0162373714559096
National Institute of Child Health and Human Development Early Child Care Research Network. (2004). Does class size in first grade relate to children's academic and social performance or observed classroom processes? *Developmental Psychology, 40*(5), 651–664. https://doi.org/10.1037/0012-1649.40.5.651
North Carolina Department of Public Instruction. (2016). Technical documentation for 2016 EVAAS analyses. Author https://ncdpi.sas.com/support/EVAAS-NC-TechnicalDocumentation-2016.pdf.
North Carolina Division of Child Development and Early Education. (2015). NC Pre-Kindergarten (NC Pre-K) program requirements and guidance. Effective SFY 2015–2016. Author www.ncchildcare.ncdhhs.gov.
North Carolina Early Learning Foundation. (2017). Shared measures of success to put North Carolina's children on a pathway to grade-level reading. Author www.buildthefoundation.org.
North Carolina Foundations Task Force. (2013). North Carolina foundations for early learning and development. Author www.ncchildcare.ncdhhs.gov.
Pearman, F. A., Springer, M., Lipsey, M., Lachowicz, M., Swain, W., & Farran, D. (2020). Teachers, schools, and pre-K effect persistence: An examination of the sustaining environment hypothesis. *Journal of Research on Educational Effectiveness, 13*(4), 1–27. https://doi.org/10.1080/19345747.2020.1749740
Peisner-Feinberg, E. S., Kuhn, L., Zadrozny, S., Foster, T., & Burchinal, M. (2020). *Kindergarten follow-up findings from a small-scale RCT study of the North Carolina Pre-Kindergarten Program*. School of Education: The University of North Carolina.
Peisner-Feinberg, E. S., LaForett, D., Schaaf, J., Hildebrandt, L., Sideris, J., & Pan, Y. (2014). *Children's outcomes and program quality in the North Carolina Pre-Kindergarten Program: 2012–2013 statewide evaluation*. FPG Child Development Institute: The University of North Carolina.
Peisner-Feinberg, E. S., Mokrova, I. L., & Anderson, T. L. (2017). *Effects of participation in the North Carolina pre-kindergarten program at the end of kindergarten: 2015–2016 statewide evaluation*. FPG Child Development Institute: The University of North Carolina.
Peisner-Feinberg, E. S., Schaaf, J., Hildebrandt, L., & LaForett, D. (2013). *Quality and characteristics of the North Carolina pre-kindergarten program: 2011–2012 statewide evaluation*. FPG Child Development Institute: The University of North Carolina.
Peisner-Feinberg, E. S., & Schaaf, J. M. (2010). Long-term effects of the North Carolina more at four pre-kindergarten program: Children's reading and math skills at third grade. In *Chapel Hill, NC*. FPG Child Development Institute: The University of North Carolina.
Peisner-Feinberg, E. S., & Schaaf, J. M. (2011). *Effects of the North Carolina more at four pre-kindergarten program on children's school readiness skills: Summary of key findings*. Chapel Hill, NC: The University of North Carolina, FPG Child Development Institute.
Peisner-Feinberg, E. S., Zadrozny, S., Kuhn, L., & Van Manen, K. (2019). *Effects of the North Carolina Pre-Kindergarten Program: Findings through Pre-K of a small-scale RCT study*. FPG Child Development Institute: The University of North Carolina.
Peugh, J. L., & Enders, C. K. (2004). Missing data in educational research: A review of reporting practices and suggestions for improvement. *Review of Educational Research, 74*(4), 525–556. https://doi.org/10.3102/00346543074004525
Phillips, D. A., Lipsey, M. W., Dodge, K. A., Haskins, R., Bassok, D., Burchinal, M. R., … Weiland, C. (2017). The current state of scientific knowledge on pre-kindergarten effects. In *Brookings*. www.brookings.edu/research/puzzling-it-out-the-current-state-of-scientific-knowledge-on-pre-kindergarten-effects.
Pianta, R. C., Downer, J., & Hamre, B. K. (2016). Quality in early education classrooms: Definitions, gaps, and systems. In S. McLanahan, J. M. Currie, R. Haskins, C. E. Rouse, & I. Sawhill (Eds.), *Starting early: Education from prekindergarten to third grade (Vol. 26, 2, pp. 119–137)* (The Future of Children).
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 https://www.acf.hhs.gov/sites/default/files/opre/head_start_report.pdf.
Raver, C. C., McCoy, D. C., Lowenstein, A. E., & Pess, R. (2013). Predicting individual differences in low-income children's executive control from early to middle childhood. *Developmental Science, 16*(3), 394–408. https://doi.org/10.1111/desc.12027
Reardon, S. F. (2019). *Educational opportunity in early and middle childhood: Variation by place and age (CEPA Working Paper No.17-12)*. Stanford Center for Education Policy Analysis. https://doi.org/10.7758/RSF.2019.5.2.03
Ronfeldt, M., Loeb, S., & Wyckoff, J. (2013). How teacher turnover harms student achievement. *American Educational Research Journal, 50*(1), 4–36. https://doi.org/10.3102/0002831212463813
Stipek, D., Clements, D., Coburn, C., Franke, M., & Farran, D. (2017). PK-3: What does it mean for instruction? *SRCD Social Policy Report, 30*(2), 1–23. https://doi.org/10.1002/j.2379-3988.2017.tb00087.x
Unterman, R., & Weiland, C. (2020). *Higher-quality elementary schools sustain the prekindergarten boost: Evidence from an exploration of variation in the Boston prekindergarten program's impacts. (EdWorkingPaper: 20-321)*. Annenberg Institute at Brown University. https://doi.org/10.26300/hkv0-wj29
Vanderhaar, J. E., Muñoz, M. A., & Rodosky, R. J. (2006). Leadership as accountability for learning: The effects of school poverty, teacher experience, previous achievement, and principal preparation programs on student achievement. *Journal of Personnel Evaluation in Education, 19*(1–2), 17–33. https://doi.org/10.1007/s11092-007-9033-8
Wechsler, D. (2014). *Technical and interpretive manual supplement: Special group validity studies with other measures and additional tables*. NCS Pearson.
Weiland, C. (2016). Impacts of the Boston prekindergarten program on the school readiness of young children with special needs. *Developmental Psychology, 52*(11), 1763–1776. https://doi.org/10.1037/dev0000168
Weiland, C., & Yoshikawa, H. (2014). Does higher peer socio-economic status predict children's language and executive function skills gains in prekindergarten? *Journal of Applied Developmental Psychology, 35*(5), 422–432. https://doi.org/10.1016/j.appdev.2014.07.001
Willoughby, M. T., Wirth, R. J., & Blair, C. B. (2011). Contributions of modern measurement theory to measuring executive function in early childhood: An empirical demonstration. *Journal of Experimental Child Psychology, 108*(3), 414–435. https://doi.org/10.1016/j.jecp.2010.04.007
Willoughby, M. T., Wylie, A. C., & Little, M. H. (2019). Testing longitudinal associations between executive function and academic achievement. *Developmental Psychology, 55*(4), 767–779. https://doi.org/10.1037/dev0000664
Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III tests of achievement*. The Riverside Publishing Company.
Zhai, F., Raver, C. C., & Jones, S. M. (2012). Academic performance of subsequent schools and impacts of early interventions: Evidence from a randomized controlled trial in head start settings. *Children and Youth Services Review, 34*(5), 946–954. https://doi.org/10.1016/j.childyouth.2012.01.026