NOT FINAL

Effects of Early Mathematics Intervention for Low-SES Pre-Kindergarten and Kindergarten Students: A Replication Study

Prentice Starkey[1], Alice Klein[1], Ben Clarke[2], Scott Baker[2], and Jaime Thomas[3]

[1] WestEd

[2] University of Oregon

[3] Format Consulting

Author Note: This article was supported by the Institute of Education Sciences, U.S. Department of Education Grant R305A120262 to P.S. The opinions expressed are those of the authors and donot represent the views of the U.S. Department of Education.

Correspondence concerning this article should be addressed to Prentice Starkey, WestEd, 2470 Mariner Square Loop, Alameda, CA 94501.

Abstract

A socioeconomic status- (SES-) related achievement gap in mathematics emerges prior to school entry, and increases as children move through elementary school. This gap makes implementation of demanding mathematics standards (e.g., the Common Core State Standards) an ongoing challenge. Early educational intervention is a strategy for addressing this challenge. A randomized controlled trial was conducted in public American preschools (1) to replicate the efficacy of an intervention, *Pre-K Mathematics*, for low-SES children, and (2) test the combined impact of this intervention and a Common-Core aligned kindergarten intervention, *Early Learning in Mathematics*. Forty-one clusters of pre-kindergarten and kindergarten classrooms, containing a sample of 389 low-SES children from an agricultural region, were randomly assigned to treatment and control conditions. The original impact findings were replicated: Child mathematics outcomes in prekindergarten were positive and significant. Gains were maintained in kindergarten. Thus, the gap can be reduced and gains maintained by sustained early intervention.

Keywords: pre-K mathematics, maths intervention, low socio-economic status, achievement gap, replication

Introduction

Many curricula used in early childhood education programs have not been evaluated or have been evaluated and found to be ineffective, relative to a business-as-usual counterfactual (Preschool Curriculum Evaluation Research Consortium, 2008). For those found to improve outcomes for 4-year-old children in pre-kindergarten (PK), research has begun to focus on the degree and persistence of these gains after children enter elementary school (Bailey, 2019). A key issue is whether early intervention sufficiently prepares low-socioeconomic (SES) children to meet demanding learning expectations for kindergarten (K; children age 5-6).

The study we report had a two-fold purpose. The first was to conduct a direct replication of an efficacy study that found that a curricular intervention, *Pre-K Mathematics*, improved mathematics outcomes for low-SES children in the PK year. Second, the study was designed to determine whether the PK gains are maintained when children subsequently receive a demanding, standards-based mathematics curriculum in K, assuming a positive result was replicated.

Gaps in Early Mathematical Knowledge

Gaps (group differences) in mathematical knowledge begin to appear in early childhood. Early mathematics gaps have been revealed through cross-national (e.g., Miller et al., 2005; Starkey & Klein, 2008) and cross-SES (e.g., Hughes, 1986; Reardon & Portilla, 2016; Starkey et al., 2004) comparisons. These differences are related to differential levels of support that children receive in their early learning environments. In the home learning environment, differences have been found in the mathematical language and the breadth and depth of mathematical concepts used in young children's home learning activities (Blevins-Knabe et al., 2000; DeFlorio & Beliakoff, 2015; Elliott & Bachman, 2017; LeFevre et al., 2010). In the preschool learning environment in the United States, public preschool programs for low-SES children have been found to be ineffective at supporting mathematics learning (e.g., U.S. Department of Health and Human Services, Administration for Children and Families, 2010). The two most widely used curricula in these programs are the Creative Curriculum and High/Scope, both of which have been evaluated and found to be ineffective in the area of mathematics in independent randomized controlled trials

(Howard, 2015; Preschool Curriculum Evaluation Research Consortium, 2008). Two studies found mixed results (that is, positive or negative, depending on the outcome) but statistically and substantively insignificant effects of the Creative Curriculum on math outcomes (Preschool Curriculum Evaluation Research Consortium, 2008). One study of High/Scope found no significant differences between treatment and control groups on child outcomes, including math outcomes (Howard, 2015; Howard & Weinberg, 2021). In summary, early gaps in mathematical knowledge are related to differences in children's learning environments. In the United States, where the most evidence has been collected, low-SES preschool children, relative to their middle-class peers, receive less support for mathematics learning both at home and in preschool. This results in a mathematics gap prior to school entry. Early intervention in PK is an approach that could potentially close this gap.

New Mathematics Standards Increase Learning Expectations

A recent structural change in mathematics education in the US, which is described below, has increased concerns about the mathematics gap prior to school entry. This change is primarily a consequence of the TIMSS (Mullis et al., 2020) and PISA (OECD, 2019) cross-national comparisons of mathematics achievement. Comparisons were then made of mathematics curricula used in the United States and curricula used in nations such as Singapore with the highest mathematics achievement (e.g., Ginsburg, et al., 2005). This led to policy recommendations to create national mathematics standards, with the objective of increasing maths skills of American students at completion of high school. These recommendations culminated in the development of the Common Core State Standards for Mathematics (CCSSM) for grades K-12 (5- to 17-year-old students) (National Governors Association Center for Best Practice & Council of Chief State

School Officers, 2010). As in high-achievement nations, the CCSSM focus on the teaching and learning of fewer mathematical topics but with more rigor, specifically with conceptually based understanding, procedural fluency, and applications at each grade level. They also provide coherence, including fine-grained sequencing of topics across grade levels.

The CCSSM is optional rather than required for state education systems, but have been adopted by 40 of the 50 states in the US. The development of higher standards for upper grades was achieved, in part, by backwards mapping from higher grades to lower grades. Consequently, when adopted, the CCSSM have increased, or in some cases established, learning expectations for students at all grade levels, beginning with kindergarten. This has created a challenge for public pre-kindergarten programs, because students now need to enter kindergarten ready for the CCSSM. To illustrate the extent of this challenge, consider two of the five CCSSM mathematical domains for kindergarten: counting and cardinality, and number and operations in base 10. Students are expected to enter kindergarten ready to learn to count to 100 by ones and tens, and to compose and decompose 11-19 into ten ones and some further ones, and to represent these actions using drawings of sets and equations. Currently, however, 41% of low-SES children finish their pre-kindergarten year in Head Start, the largest public preschool program in the US, unable to count sets of 10 objects correctly (U.S. Department of Health and Human Services, Administration for Children and Families, 2010). Thus, an early mathematics gap makes it difficult for schools to begin moving a substantial number of students through the CCSSM at their grade level.

The Original Efficacy Study of *Pre-K Mathematics*

The original study was conducted as part of the Preschool Curriculum Evaluation Research (PCER) program of the Institute of Education Sciences, United States Department of Education

(Klein et al, 2008). Teachers implemented *Pre-K Mathematics* (Klein & Starkey, 2004) in public preschool programs, including federally funded Head Start programs and state-funded programs in California and New York in the United States. Randomisation was at the class level where teachers' classrooms were randomly assigned to treatment and business-as-usual control conditions. The sample was comprised of low-SES children who were age-eligible for K in the following school year. Children's mathematical knowledge was assessed using the Child Math Assessment (CMA) in fall and spring of PK (see Child Measures, below, for a description of the CMA). Implementation was measured by multiple metrics, including fidelity (see Table 1). The same fidelity instrument, the Fidelity of Implementation Record Sheet, was used in both this study and the original study (see the Appendix of Klein et al., 2008). The *Pre-K Mathematics* intervention was found to be effective. Math outcomes at the end of PK, as measured by the CMA, were statistically significantly greater for treatment children than for control children ($p < .0001$; Cohen's d ES = 0.55; 95% CI: 0.330, 0.769).

- Insert Table 1 here –

Comparison of the Original and Replication Studies

A direct replication was conducted. The original and replication studies were identical or very similar in regard to several key features:

Experimental Design. A 2- arm randomized controlled trial (RCT), with Pre-K Mathematics treatment condition and business-as-usual control, was used in both studies.

Intervention. Both studies evaluated the Pre-K Mathematics curriculum, using the same implementation metrics (see Intervention Fidelity and Classroom and Home Curriculum Dosage, Table 1).

Teacher Training. The same training model was used, including Pre-K Mathematics workshops (see Teacher Workshops, Table 1) and in-classroom curriculum coaching (see Coaching Visits per Classroom, Table 1).

Student Assessment. Pretest and posttest assessments of children's mathematical knowledge were conducted using the same instrument, the Child Math Assessment.

Study Participants-Teachers. Only public PK teachers participated; teaching experience was similar in both studies (Table 2)

Study Participants-Students. Both samples were comprised of low-SES 4-year-olds.

Education Programs. Only public PK programs (state preschools and federal Head Start) were included in both studies.

Some differences between the studies bear on the external validity of the original study:

Study Participants-Teachers. Teachers had less education in the replication study (Table 2).

Study Participants-Students. Ethnic and racial composition of children differed (Table 3).

Education Programs. Programs were situated in urban areas in the original study and in agricultural areas in the replication study. This study can therefore be considered a close rather than a direct replication.

- Insert Tables 2 and 3 here –

Moderation of Early Mathematics Outcomes by Children's Executive Attention

A secondary purpose of the replication study was to identify subgroups of children for whom curricular intervention is less effective. Of those interventions found to be generally effective for the intended population, few have also been evaluated to determine whether they are less effective for specific subgroups of children (Fuchs & Fuchs, 2019). A growing body of evidence suggests that a complex relation exists between children's developing mathematical knowledge and executive attention (e.g., Cueli, et al., 2020; Peterson et al., 2017). Teacher ratings of inattention by elementary school children predict mathematics achievement and skills (Cirino et al., 2007; Fuchs et al., 2006; Raghubar, et al., 2009). Also, executive attention abilities of preschool children predict subsequent mathematical knowledge (Blair & Razza, 2007; Kroesbergen et al., 2009). We have obtained evidence for a relation between mathematical knowledge and executive attention in low-SES PK children in a recent intervention project (Lonigan et al., 2015). Specifically, attention focusing and inhibitory control, as measured by the Child Behavior Questionnaire (CBQ) at the beginning of PK, predicted mathematical knowledge at the end of PK and K (Barnes et al., 2021). Given these findings, we included the CBQ in the replication study to further examine relations between executive attention and mathematical knowledge in young children.

Research Questions

Confirmatory Research Question

Will the *Pre-K Mathematics* intervention lead to increased mathematical knowledge in treatment children, relative to control children, by the end of PK?

Exploratory Research Questions

Will the gains in PK be maintained in K when treatment children receive a mathematics curriculum, *Early Learning in Mathematics* (*ELM*), aligned with demanding mathematics standards?

Is the effectiveness of the intervention moderated by children's executive attention in PK or K?

Method

Design

The principal effects of interest are increases in mathematical knowledge of children who receive a mathematics curriculum in PK and a standards-based mathematics curriculum in K. The experimental design was a cluster randomized controlled trial (RCT) in which 41 clusters (20 treatment and 21 control) of PK and K classrooms were randomly assigned to the treatment condition (*Pre-K Mathematics* and *Early Learning in Mathematics*) or the business-as-usual control condition (existing classroom curriculum). These clusters were enrollment patterns where children from a PK classroom were slated to attend 1-2 in-cluster K classrooms the following year. Children from a PK classroom in a cluster were enrolled in 1-2 K classrooms (37 treatment and 27 control) in the cluster. In total 41 PK and 64 K classrooms were included.

Prior to randomization of the classroom clusters, all children in the research sample were identified. Thus, no children joined the sample after randomization. Randomization was performed within school districts, and teachers in designated clusters gave their consent at the beginning of the PK year to ensure compliance with randomization procedures.

In PK, 11 of the business-as-usual C classrooms used a comprehensive curriculum (Creative Curriculum, Frog Street, or Houghton Mifflin) with a mathematics component, and 10 had no mathematics curriculum. In K, all of the business-as-usual classrooms had mathematics curricula, and 9 out of 27 were in school districts that had adopted at least one CCSSM-aligned mathematics curriculum for grade K. The remainder used unaligned curricula, including older curriculum editions by major publishers. The What Works Clearinghouse has not rated as effective any of the PK or K mathematics curricula used in the control classrooms. Teachers in PK and K control classrooms received no training in the RCT pertaining to the new pre-K maths curriculum.

Participants

Programs and Teachers

The study was conducted in public, federally funded Head Start programs and in California state-funded preschool programs and elementary schools in predominantly agricultural regions of California in the United States. The majority of families served by the participating programs were Hispanic/Latino.

Head Start and California state preschool programs are categorical rather than universal, requiring low income for families to qualify. These programs enroll 3- and 4-year-old children. The elementary schools were from 10 small school districts. Their kindergarten classrooms enroll 5-year-old children. The highest education level and amount of teaching experience of teachers in the original study (Klein et al., 2008) and replication study are given in Table 2.

Children

Children who did not have a significant neurodevelopmental disorder and who were age-eligible for entrance into public kindergarten the following year were randomly selected to participate in the study. Within each classroom, 8-12 children (*M* = 9.5), half boys and half girls, when possible, were randomly selected. Teachers knew which children were pretested and, therefore, which children were included in the study.

The sample included 389 children (197 treatment and 192 control) from low-income families. Demographic characteristics of children at baseline in the original study and replication study are given in Table 3. Children's age at baseline assessment was 4.5 years (treatment, *M* = 4.5; control, *M* = 4.5; range: 4.0 – 5.0 years). Girls comprised 55% of children in the treatment condition and 53% in the control condition. At pretest, 48% of the children were assessed in English (treatment, 51%; control, 44%), 41% in Spanish (38% treatment; 43% control, and 12% bilingually (treatment, 11%; control, 13%).

Attrition. The rate of measurement attrition across the PK year was 4.6% overall and 1.1% differential for the CMA, and 4.4% overall and 0.6% differential for the Test of Early Mathematics Ability (TEMA-3). The What Works Clearinghouse guidelines place this rate of attrition within the green rating zone, labeled as *low attrition*, with a tolerable threat of bias (What Works Clearinghouse, 2020, p. 10). There was no attrition at the classroom level.

Ethics Approvals. The WestEd Institutional Review Board (IRB) reviewed and approved the human subjects protocol for this study (Protocol ID 201201-02). Approved informed-consent forms to participate in the study were administered to, and signed by, all participating teachers and by a parent or legal guardian of all participating children, prior to the onset of data collection (IRB approval letter available upon request).

The Pre-Kindergarten Intervention: Pre-K Mathematics

A curricular intervention, *Pre-K Mathematics*, was implemented in the PK year in treatment classrooms for the purpose of mathematically enriching children's preschool and home learning environments. Treatment teachers added this curriculum to their existing set of classroom curricula. Teachers who had implemented a whole child curriculum with a mathematics component, or who had supported mathematics learning through use of eclectic instructional practices, were allowed to continue doing so. Control teachers were directed not to alter their curricular choices or instructional practices from business as usual.

Pre-K Mathematics is comprised of classroom and home mathematics activities, and it is implemented according to a standard curriculum plan. Additionally, teachers enrich the classroom by installing a mathematics learning center and utilizing mathematics software or apps. Teachers encourage children to use the mathematics learning center and other classroom resources during learning center periods.

The classroom mathematics activities included concrete manipulatives and were conducted by teachers with children in small groups. The home mathematics activities included picture strips as guides, and were conducted by parents (or other adults in the home) with children in parent-child dyads. A teacher's manual provided the curriculum plan, which linked a sequence of small-group classroom activities to home activities, and an implementation schedule with specified time for teachers to review small-group mathematics activities with children who had been absent or who were continuing to experience difficulty with a particular activity. Curriculum units, consisting of sets of mathematically related activities, were (1) Number and Number Relations, (2) Arithmetic Operations (Fall Activities), (3) Spatial Sense and Geometry, (4) Patterns, (5) Arithmetic Operations (Spring Activities), and (6) Measurement and Data. Classroom mathematics activities were conducted with small groups of children (typically 4 children per group) twice during the week for approximately 15-20 minutes per group.

Curriculum dosage was tracked in PK and found to average 85% of the recommended number of small group activity sessions per child. This average was below the maximum possible primarily due to doses missed during child absences, which teachers did not subsequently make up. Children's performance on individual activities during small group sessions were assessed by teachers. Teachers gave a rating of mastery on an activity if the child successfully engaged in independent problem solving during the entire activity (i.e., made no errors and received no scaffolding). Teachers were trained to continue providing dosage (up to three additional doses) for non-mastered activities to individual children, as needed and as instructional time permitted, until mastery was achieved. According to teachers' end-of-year records, the mean proportion of small-group mathematics activities that children mastered was 0.59 (*SD* = 0.25; range, 0.08-1.00. For comparison purposes, key implementation metrics for the original *Pre-K Mathematics* efficacy study and the present replication study are given in Table 1.

Teachers sent *Pre-K Mathematics* home activities to parents for use at home over the course of the year. Materials and instructions were provided in English and Spanish. Teachers asked parents to return a Parent Feedback Form after completing each activity. Parental reports indicated that the mean proportion of home activities used by families used was 0.94 (*SD* = 0.17; range, 0.27-1.00).

Treatment teachers were trained through six days (45 hours) of mandatory *Pre-K Mathematics* curriculum workshops, and through on-site coaching provided in each classroom by project trainers 1-2 times per month (*M* = 12.0; *SD* = 2.51). Teachers learned the *Pre-K Mathematics* curriculum, mathematics instructional practices (mathematically focused small-group activities), how to use of implementation tools (progress monitoring and dosage tracking), and strategies to support parents' use of home mathematics activities. Teachers' intervention fidelity scores were high (*M* = 0.97; *SD* = 0.04; range, 0.84-1.00). Thus, the curriculum was generally implemented as intended by the developers.

The Kindergarten Intervention: *Early Learning in Mathematics*

A curricular intervention, *Early Learning in Mathematics (ELM)*, was implemented in the kindergarten year in treatment classrooms. *ELM* (Chard et. al., 2008) is a whole-class mathematics curriculum that is designed for the general population of K students. It is not specifically designed for at-risk students. The content strands of *ELM* are (1) whole numbers and operations, (2) geometry, and (3) measurement. These strands align directly with the CCSSM for kindergarten. It consists of 120 mathematics lessons, 45 minutes in duration, supplemented by a daily whole-class calendar activity lasting 15-minutes.This activity includes a calendar routine (e.g., days of the week, dates on the monthly calendar; days of the year using a place value chart), with learning objectives that change quarterly.. All *ELM* activities incorporate a variety of mathematics models for children to build conceptual understanding. The curriculum also includes activities that are sent home weekly, in English and Spanish, for parents to provide their children with additional mathematics practice outside of school.

ELM has been evaluated in an Institute of Education Sciences-funded efficacy study (Clarke et al., 2011). Teacher's classrooms were randomly assigned to treatment and control conditions. The sample was comprised of K children from low- and middle-SES families. Children's mathematical knowledge was assessed using the TEMA-3 in fall and spring of K (see Child Measures, below, for a description of the TEMA-3). Implementation fidelity was measured on dosage, instruction fidelity, number of teacher workshops and coaching sessions. *ELM* was found to be effective, as mathematics outcomes at the end of K, as measured by the TEMA-3, were significantly greater for treatment children than for control children (p = .0017; Hedges's g ES = 0.24; 95% CI = 0.10, 0.39).

For the present study, *ELM* was adapted to align closely with the demanding grade K standards contained in the CCSSM. K treatment teachers used this curriculum as their regular classroom mathematics curriculum. Control teachers were directed not to alter their mathematics practices or curriculum from business as usual.

Treatment teachers were trained through four days (30 hours) of mandatory *Early Learning in Mathematics* curriculum workshops, and through on-site coaching provided in each classroom by project trainers 1-2 times per month (M=12.0; SD=0.78). At the initial *ELM* workshop, teachers received all curricular materials, including quarterly based teacher guides and classroom mathematics kits. They were given a brief overview of the *ELM* curriculum, its instructional objectives, the critical content of kindergarten mathematics, and the instructional practices that have been empirically validated to increase student mathematics achievement (e.g., teacher-provided academic feedback). Workshops focused on implementation fidelity to the *ELM* curriculum and the delivery of effective kindergarten mathematics instruction. Teachers had an opportunity to practice with sample activities from the curriculum and received feedback on their instructional delivery from project trainers. Teachers' instruction fidelity scores (the degree to which teachers adhere to the activities of the ELM lesson) were, on average, high (M = 0.92; SD = 0.07; range, 0.70-1.00). Thus, the curriculum was generally implemented as intended by the developers.

Curricula in Control Classrooms and Grade 1

Control teachers were provided with no training before or during the school year in which the RCT was conducted. They continued to provide their usual mathematics instruction. Information about the mathematics curriculum and mathematics instruction used in control classrooms was obtained from a teacher questionnaire and from public information provided by school districts regarding textbook adoptions. These showed that 38% of PK control teachers used a comprehensive curriculum that contained a mathematics component. None of these curricula has been given a positive effectiveness rating for mathematics by the What Works Clearinghouse. Some, such as the Creative Curriculum, have been evaluated and assigned an effectiveness rating of zero (not effective, relative to other curricula). Others have not been evaluated. The remainder of the teachers used no mathematics curriculum. Unlike Pre-K Mathematics, the Business-as-Usual (BAU) curricula did not include intentional, small-group mathematics activities, and teachers had received little curriculum-focused training in mathematics. Thirty-three percent of the K control teachers used a CCSSM-aligned mathematics curriculum. No control teachers used K mathematics curricula that had been rated as effective by the What Works Clearinghouse. District-provided information indicated that a similar percentage of grade 1 teachers used a CCSSM-aligned mathematics curriculum, and none used mathematics curricula that had been rated as effective.

Child Measures and Data Collection Procedures

Two measures of children's mathematical knowledge were used to determine the impact of the intervention. In addition, the Child Behavior Questionnaire (CBQ; Rothbart et al., 2001), which measures children's attention focusing, inhibitory control and impulsivity, was used to examine moderation of impacts of the intervention.

Children's mathematical knowledge was assessed by two measures, the Child Math Assessment (CMA; Milburn et al., 2019) and the Test of Early Mathematics Ability, 3rd Edition (TEMA-3; Ginsburg & Baroody, 2003). The CMA measures preschool children's informal mathematical knowledge across the broad range of concepts and skills that research has shown are developing at this age, including number, arithmetic operations, space and geometry, informal measurement, and patterns. The CMA is appropriate for children from 3 to 5 years of age. It is comprised of 9 tasks, with multiple problems per task, and individual problems on each task are scored for accuracy (0/1). Then a mean composite score is obtained by averaging across tasks. Test-retest reliability of the CMA is 0.91, and internal consistency (stratified coefficient alpha) is 0.92. Concurrent validity with the TEMA-3 is 0.74 ($p < .01$).

The TEMA-3 measures informal and formal mathematical knowledge. It is a standardized instrument that is appropriate for children from 3 to 8 years of age. The TEMA provides an in-depth assessment of children's numerical abilities, including items that measure number sense, knowledge of numerals, arithmetic problem solving and calculation skills. Each item that a child receives is scored for accuracy (0/1), and then a total score is obtained based in the number of correct items. Since the CMA measures all of the conceptual dimensions comprising children's early mathematical knowledge, it was better aligned with the *Pre-K Mathematics* intervention and served as the primary outcome measure for the PK year. The TEMA, however, can be used with a wider age range, including children in PK, K and Grade 1. Thus, the TEMA was included in this study to measure math outcomes in K and Grade 1. Test-retest reliability of the TEMA-3 ranges from 0.82 to 0.93, and internal consistency (stratified coefficient alpha) is 0.92.

Assessors administered the two measures in separate sessions, with the orders in which these measures were administrered being counterbalanced in both the treatment and control conditions. All assessors attended training sessions to learn the CMA and TEMA-3, and then underwent a rigorous certification procedure to ensure that they administered the measures with a high level of proficiency. Assessors were blind to which condition children had been placed (i.e., treatment, control). Children were assessed in their primary language (English or Spanish) or bilingually, as needed, with conceptual scoring. Assessments were conducted in a quiet room at children's schools. The CMA was administered individually to children in one 20-minute session, and the TEMA was administered in a 30-minute session. The pretest assessment was conducted during the first six weeks of the PK school year prior to implementation of the intervention. Posttest and follow-up assessments were conducted during the last six weeks of the school year, after the PK intervention was completed, and in the last six weeks of grades K and 1. All data were scored independently by two research assistants, and any discrepancies were resolved before data were entered into a database. Data sets were entered independently by two research assistants, and any discrepancies were resolved before the database was finalized.

The Child Behavior Questionnaire (Rothbart, et al., 2001), as adapted for use with teachers (Eisenberg, et al., 2004), consists of three subscales: attention focusing, inhibitory control, and impulsivity. For each subscale, the classroom teacher rates a series of statements about the child's behavior. The CBQ has been found to be a highly reliable measure of children's effortful control behavior, defined as the "efficiency of executive attention" (Rothbart & Bates, 2006) Cronbach's alpha (also referred to as coefficient alpha), which measures the internal consistency or reliability of a set of items within a test, has been calculated for the CBQ by Eisenberg, et al. (2004). Values above 0.80 are considered to be very robust. For the attention focusing subscale, alpha =0.85; for the inhibitory control subscale, alpha = 0.88; for the impulsivity subscale, alpha

= 0.88. All participating PK and K treatment and control teachers completed the CBQ for children in their classroom who were in the research sample. These data were collected in winter (mid-year) of the school year.

To recapitulate, this new study replicated the original study in its implementation and delivery dosage (see Table 1 for the implementation metrics), but differed slightly in the teacher educational level (in the replicated study teachers had less education) and ethnic and racial compostion of the students. While the original study was conducted in urban schools, this replication was carried out in schools in agricultural areas.

Results

Child Mathematics Outcomes in the Pre-Kindergarten Year

The first set of analyses addressed the confirmatory research question, which was the focus of this replication: Will the *Pre-K Mathematics* intervention lead to increased mathematical knowledge in treatment children, relative to control children, by the end of PK?

As justified above, the CMA was the principal measure of child mathematics outcomes in PK, and the TEMA-3 was used to make follow-up assessment in elementary school possible. Treatment and control group scores on the CMA and TEMA-3 at PK pretest and posttest are given in Tables 4 and 5. Equations and values used to calculate baseline equivalence and impacts are provided in Appendix A.

Children in the treatment and control groups had mathematics scores that were equivalent at baseline, both for the CMA (p = .774) and the TEMA-3 (p = .670). Intent-to-treat (ITT) effects of the *Pre-K Mathematics* intervention on outcomes measured at the end of PK were estimated to answer the study's confirmatory research question using the model specification described in

Equation 1 (see Appendix A). The analysis found a statistically significant positive impact of the *Pre-K Mathematics* intervention at the end of PK for the CMA ($p < .0001$; Cohen's d ES = 0.52; 95% CI = 0.309, 0.730; see Equations 2 and 3 of Appendix A, and Table 5). A statistically significant positive impact was also found for the TEMA-3 ($p = .003$; Cohen's d ES = 0.30; 95% CI = 0.094, 0.505).

- Insert Tables 4 and 5 here –

Child Mathematical Knowledge in Kindergarten and Grade 1

The next set of analyses addressed the first exploratory research question: Will the gains in PK be maintained in K when treatment children receive a mathematics curriculum, ELM, aligned with demanding mathematics standards?

The TEMA-3 was used to assess longitudinal outcomes for the treatment-on-the-treated (TOT) sample. Prior to data analysis, the developers of *Pre-K Mathematics* and *ELM* set a criterion for inclusion in TOT analyses. Treatment children must receive at least 75% of the intended treatment in PK and 75% in K, and control children must attend for at least 75% of the PK school year and 75% in K, to be considered to have received the intended curriculum.

The analysis used a cross-classified, random-effects, longitudinal growth curve model that included TEMA-3 scores at four waves of data collection (fall of PK, and spring of PK, K and 1) as well as PK and K classroom IDs. The repeated measures portion of the model was estimated using an unstructured variance-covariance matrix, since variability of the TEMA-3 increased over the four waves. The between-subjects portion of the model included condition, wave, and wave by condition, with gender and pretest age as covariates. The longitudinal analysis found that the impact at the end of K was positive and statistically significant, indicating that the mathematics intervention did, in fact, have an impact on mathematical knowledge in treatment children, relative to control children, in the K year (Table 6). The impact at the end of grade 1 was not statistically significant (Table 6).

Pairwise comparison of the wave by condition coefficients at waves 2 and 3 were conducted to determine whether the size of the impact at the end of PK (wave 2) was maintained in grade K (wave 3). These comparisons showed that the wave by condition coefficients at waves 2 and 3 were not significantly different from one another at conventional levels ($p = .098$). Thus, the null hypothesis that the impacts are equal at waves 2 and 3 cannot be rejected.

- Insert Table 6 here –

Moderation of the Effectiveness of the Intervention by Executive Attention

The next set of analyses addressed the second exploratory research question: Is the effectiveness of the intervention moderated by children's executive attention in PK or K? Moderation analysis for TEMA-3 was conducted using the same method as the confirmatory analysis of impacts at the end of the PK and K years described above, with the addition of moderator variables. The potential moderators were included, one at a time, with the interaction of the moderator with condition (moderating the level differences). The latter term investigated how CBQ scores moderated the condition effect. Models were run with continuous TEMA-3 raw scores. Detailed results are given for those cases where a statistically significant or substantively important moderation effect was detected. When there was a statistically significant or substantively important moderation effect for the continuous moderation variable, we unpacked this effect by categorizing children's scores on the CBQ as low (approximately 1 *SD* below *M*) or high (1 *SD* above *M*) in PK and again in K, and examining differences in TEMA-3 raw scores between treatment and control children at each level.

Scores on each CBQ subscale (attention focusing, impulsivity, and inhibitory control) in the PK year were examined separately. No evidence of moderation was found for impulsivity or attention focusing. Specifically, for these two subscales, the coefficient on the moderator by condition interaction term was small in magnitude and not statistically significantly different from zero. Inhibitory control, however, did moderate the effect of condition (coefficient = 3.3, $p = 0.04$). When inhibitory control scores were low (approximately 1 *SD* below *M*), the effect size of the difference between TEMA-3 raw scores for treatment and control children was small (Cohen's d ES = 0.13). When inhibitory control scores were high (approximately 1 *SD* above *M*), the effect size of the difference in TEMA-3 raw scores was moderate (Cohen's d ES = 0.52).

Scores on each CBQ subscale in the K year were also examined separately. No evidence of moderation was found for impulsivity. Attention focusing strongly moderated the effect of condition (coefficient = 4.0, $p = 0.01$). When attention focusing scores were low, the effect size of the difference between TEMA-3 raw scores for treatment and control children in spring K was very small (Cohen's *d* ES = 0.02), but when scores were high, the effect size was large (Cohen's *d* ES = 0.67). A similar, though less pronounced, pattern was found for inhibitory control scores in the K year (coefficient on the moderator by condition interaction term = 3.9, $p = 0.09$). When scores were low, treatment children outperformed control children by a smaller margin (Cohen's *d* ES = 0.12) than when scores were high (Cohen's *d* ES = 0.58).

Discussion

Replication of Findings of the Original Efficacy Study

Implementation metrics of the original and replication studies indicated that implementation of *Pre-K Mathematics* was conducted similarly (Table 1). Teachers received similar training (workshops and in-classroom coaching), and this training enabled teachers to teach with adequate (.80) to high (.90+) fidelity and dosage delivery. Thus, *Pre-K Mathematics* was implemented as intended.

The main child outcomes were replicated. Both the original efficacy study and the replication study found significant, positive child outcomes in mathematics at the end of PK. Effect sizes on the Child Math Assessment measure were similar in the original study ($p < .0001$; Cohen's d ES = 0.55; 95% CI: 0.330, 0.769) and the replication study ($p < .0001$; Cohen's d ES = 0.52; 95% CI = 0.309, 0.730). Together these findings constitute a strong body of evidence for the efficacy of *Pre-K Mathematics*.

Maintenance of Early Gains

This study was also conducted to determine whether children's statistically significant math gains in PK were maintained in K when children subsequently received *ELM*, a demanding mathematics curriculum. Recall that this curriculum was aligned to meet the high learning expectations recommended by the developers of the Common Core State Standards for Mathematics. Implementation metrics indicated that K teachers received the recommended training (workshops and in-classroom coaching), and then implemented with a high degree of instructional fidelity. Thus, *ELM* was implemented as intended.

Child mathematics impacts were statistically significant at the end of K. Furthermore, pairwise comparisons showed that the wave by condition coefficients at waves 2 (end of PK) and 3 (end of K) were not statistically significantly different from one another. These values were similar, therefore the relative gains treatment children made in PK did not fade, and instead, were maintained in K. In grade 1, instruction again reverted mostly to use of curricula that were unaligned with the CCSSM or that had not been evaluated. In this context, gains were no longer fully maintained. To summarize, the mathematics learning of low-SES children across the PK and K years was accelerated. It can be concluded that the early SES-related mathematics gap can be reduced through high quality mathematics instruction in the PK and K years.

Moderation of Mathematics Gains by Executive Attention

Although the intervention was effective for most children in the sample, it was not equally effective for some sub-groups. Components of children's executive attention moderated the effectiveness of the intervention. The intervention was less effective for PK children with low inhibitory control. The intervention(s) were less effective for K children with low inhibitory control and K children with low attention focusing. This pattern of findings suggests that the importance of executive attention for mathematical learning and development may increase from PK to K. Further research is needed to understand why this pattern of moderation occurred. For example, there is a need to disentangle children's age (4 in PK and 5 in K) and the instructional setting for mathematics (small group instruction in PK and whole group instruction in K). Research could help identify of sub-groups of young children who are at risk for mathematical learning difficulties, and it could lead to a better understanding of cognitive factors that put them at risk.

Limitations of This Work and Next Steps

Based on the prior evidence that *ELM* alone is effective (Clarke et al., 2011), it can be argued that the maintenance of PK gains in K may be due, to an unknown degree, to the efficacy of *ELM*. Further research could determine the individual contributions of *Pre-K Mathematics* and *ELM*, and the possible synergistic contribution of yoking together effective PK and K curricula. Also, it is noteworthy that mathematics impacts lessened in grade 1. At this grade level, approximately two-thirds of teachers used mathematics curricula that were not closely aligned to the CCSSM and none used curricula that had not been shown to be effective. Further research could determine whether the use of an effective high-standards curriculum in grade 1 can maintain or strengthen relative gains present at the end of K. A limitation of the exploratory research on executive attention is the reliance on teacher ratings of child behavior rather that direct assessments. Research using both types of measures is needed.

Now that the efficacy of the Pre-K Mathematics intervention has been replicated, we think three next steps are warranted. First, a study of the long-term sustainability of implementation by public PK programs is needed. We have recently completed a sustainability study and have reported preliminary findings (Starkey, Flynn, & Klein, 2021, September 28). Second, an effectiveness study is needed to determine whether public PK programs can implement effectively while providing at least some of the resources, such as staff to provide on-site curriculum coaching, that our efficacy grant provided. A third step is to determine whether Pre-K Mathematics is scalable. A scaling study would provide important evidence about the external validity of the intervention.

In conclusion, this study confirms that an enriching early mathematical learning environment is important for the development of mathematical skills for young children. Therefore, future work in this area to understand how best to support early mathematical learning and development in all children will be particularly valuable. Young children in many parts of the world would no doubt benefit from enrichment of their mathematical learning environment.

References

Bailey, D.H. (2019). Explanations and implications of diminishing intervention impacts across time. In D. Berch & D.C. Geary (Eds.), *Mathematical cognition and learning*, Vol. 5, pp. 321-346.

Barnes, M. A., Klein, A., Starkey, P., Eisenberg, N., & Spinard, T. (2021). Relations of attention and mathematical knowledge from pre-kindergarten to grade 1. [Manuscript in preparation]. Psychological Sciences, Vanderbilt University.

Blair, C., & Razza, R. P. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Development, 78*, 647-663.

Blevins-Knabe, B., Austin, A., Musun, L., Eddy, A., & Jones, R. (2000). Family home care providers' and parents' beliefs and practices concerning mathematics with young children. *Early Child Development and Care*, *165*, 41-58.

Chard, D., Baker, S. K., Clarke, B., Jungjohann, K., Davis, K. L. S., & Smolkowski, K. (2008). Preventing early mathematics difficulties: The feasibility of a rigorous kindergarten mathematics curriculum. *Learning Disabilities Quarterly, 31*(1), 11-20.

Cirino, P. T., Fletcher, J. M., Ewing-Cobbs, L., Barnes, M. A., & Fuchs, L. S. (2007). Cognitive arithmetic differences in learning difficulty groups and the role of behavioral inattention. *Learning Disabilities Research & Practice, 22*, 25 – 35.

Clarke, B., Smolkowski, K., Baker, S., Fien, H., Doabler, C., & Chard, D. (2011). The impact of a comprehensive tier 1 core kindergarten program on the achievement of students at-risk in mathematics. *Elementary School Journal*, *111*, 561-584.

Cueli, M., Areces, D., García, T., Alves, R.A., & González-Castro, P. (2020). Attention, inhibitory control and early mathematical skills in preschool students. *Psicothema*, *32*, 237-244.

DeFlorio, L., & Beliakoff, A. (2015). Socioeconomic status and preschoolers' mathematical knowledge: The contribution of home activities and parent beliefs. *Early Education and Development*, *26*(3), 319–341. https://doi.org/10.1080/10409289.2015.968239

Eisenberg, N., Spinrad, T.L., Fabes, R.A., Reiser, M., Cumberland, A., Shepard, S.A., Valiente, C., Losoya, S.H., Guthrie, I.K., Thompson, M., & Murphy, B. (2004). The relations of effortful control and impulsivity to children's resiliency and adjustment. *Child Development*, *75*, 25-46.

Elliott, L. & Bachman, H.J. (2017). How do parents foster young children's math skills? *Child Development Perspectives*, *12*, 16-21.

Fuchs, D. & Fuchs, L.S. (2019). On the importance of moderator analysis in intervention research: An introduction to the special issue. *Exceptional Children*, *85*, 126-128.

Fuchs, L. S., Fuchs, D., Compton, D. L., Powell, S. R., Seethaler, P. M., Capizzi, A. M. et al. (2006). The cognitive correlates of third-grade skill in arithmetic, algorithmic computation, and arithmetic word problems. *Journal of Educational Psychology, 98*, 29-43.

Ginsburg, A., Leinwand, S., Anstrom, T., & Pollock, E. (2005). What the United States Can Learn From Singapore's World-Class Mathematics System. https://files.eric.ed.gov/fultext/ED491632.pdf

Ginsburg, H., & Baroody, A. J. (2003). *TEMA-3: Test of early mathematics ability*. Pro-Ed.

Howard, E. (2015). An efficacy trial of the High Scope preschool curriculum. Abstract for IES Grant R305A150049 https://ies.ed.gov/funding/grantsearch/details.asp?ID=1717

Howard, E. & Weinberg, E. (2021). High Scope preschool curriculum and professional development efficacy study. AIR Project Report. https://www.air.org/project/highscope-preschool-curriculum-and-professional-development-efficacy-study

Hughes, M. (1986). *Children and number: Difficulties in learning mathematics*. Oxford: Blackwell.

Klein, A., & Starkey, P. (2004). *Scott Foresman – Addison Wesley Mathematics: Pre-K*. Glenview, IL: Pearson Scott Foresman.

Klein, A., Starkey, P., Clements, D., Sarama, J., & Iyer, R. (2008). Effects of a pre-kindergarten mathematics intervention: A randomized experiment. *Journal of Research on Educational Effectiveness*, *1*(3), 155–178. https://doi.org/10.1080/19345740802114533

Kroesbergen, E. H., Van Luit, J. E. H., Van Lieshout, E. C. D. M., Van Loosbroek, E., & Van de Rijt, B. A. M. (2009). Individual differences in early numeracy: The role of executive functions and subitizing. *Journal of Psychoeducational Assessment, 27*, 226-236.

LeFevre, J., Polyzoi, E., Skwarchuk, S., Fast, L., & Sowinski, C. (2010). Do home numeracy and literacy practices of Greek and Canadian parents predict the numeracy skills of kindergarten children? *International Journal of Early Years Education*, *18*(1), 55–70. https://doi.org/10.1080/09669761003693926

Lonigan, C.J., Phillips, B.M., Clancy, J.L., Landry, S.H., Swank, P.R., Assel, M., Taylor, H.B., Klein, A., Starkey, P., Domitrovich, C.E., Eisenberg, N., de Villers, J., de Villers, P., Barnes, M. A., and the School Readiness Consortium (2015). Impacts of a comprehensive school readiness curriculum for preschool children at risk for educational difficulties. *Child Development*, *86* (6), 1773-1793.

Milburn, T. F., Lonigan, C. J., DeFlorio, L., & Klein, A. (2019). Dimensionality of preschoolers' informal mathematical abilities. *Early Childhood Research Quarterly*, *4* (2), 487-495. https://doi.org/10.1016/j.ecresq.2018.07.006

Miller, K. F., Kelly, M., & Zhou, X. (2005). Learning mathematics in China and the United States. In J. I. D. Campbell (Ed.), *Handbook of mathematical cognition* (163–178). New York: Psychology Press.

Mullis, I.V.S., Martin, M.O., Foy, P., Kelly, D.L., & Fishbein, B. (2020). *TIMSS 2019 international results in mathematics and science*. https://timss2019.org/reports/

National Governors Association Center for Best Practice & Council of Chief State School Officers. (2010). *Common core state standards for mathematics*. http://www.corestandards.org/assets/CCSSI_Math Standards.pdf

OECD (2019) *PISA 2018 Results (Volume I): What students know and can do*. PISA, OECD Publishing, Paris, https://doi.org/10.1787/5f07c754-en.

Peterson, R.L., Boada, R., McGrath, L.M., Willcutt, E.G., Olson, R.K., & Pennington, B.F. (2017). Cognitive prediction of reading, math, and attention: shared and unique influences. *Journal of Learning Disabilities*, *50*, 408-421.

Preschool Curriculum Evaluation Research Consortium. (2008). *Effects of preschool curriculum programs on school readiness (NCER 2008-2009)*. National Center for Education Research, Institute of Education Sciences, U.S. Department of Education.

Raghubar, K. P., Cirino, P., Barnes, M. A., Ewing-Cobbs, L., Fletcher, J., & Fuchs, L. (2009). Errors in multi-digit arithmetic and behavioral inattention in children with math difficulties. *Journal of Learning Disabilities, 42*, 356-371.

Reardon & Portilla (2016). Recent trends in income, racial, and ethnic school readiness gaps at kindergarten entry. *AERA Open*, *2(3)*, 1-18. https://journals.sagepub.com/doi/pdf/10.1177/2332858416657343

Rothbart, M.K., Ahadi, S.A., Hershey, K., & Fisher, P. (2001). Investigations of temperament at three to seven years: The Children's Behavior Questionnaire. *Child Development*, 72, 1394-1408.

Rothbart, M.K. and Bates, J.E. (2006) Temperament. In W. Damon, & N. Eisenberg (Eds.) *Handbook of Child Psychology: Volume 3, Social, Emotional, and Personality Development*, 6th Edition (pp. 105-176). Wiley, New York.

Starkey, P., Flynn, K., & Klein, A. (2021, September 28). Continuous Improvement of a What Works Clearinghouse Rated Early Mathematics Intervention [Paper Presentation]. Society for Research on Educational Effectiveness 2021 Conference, Arlington, VA, United States.

Starkey, P., & Klein, A. (2008). Sociocultural influences on young children's mathematical knowledge. In O.N. Saracho & B. Spodek (Eds.) *Contemporary perspectives on mathematics in early childhood education*. (pp.253-276). Charlotte, NC: Information Age Publishing.

Starkey, P., Klein, A., & Wakeley, A. (2004). Enhancing young children's mathematical knowledge through a pre-kindergarten mathematics intervention. *Early Childhood Research Quarterly, 19,* 99-120. https://doi.org/10.1016/j.ecresq.2004.01.002

U.S. Department of Health and Human Services, Administration for Children and Families (January 2010). *Head Start impact study. final report*. Washington, DC. https://www.acf.hhs.gov/sites/default/files/documents/opre/hs_impact_study_final.pdf

What Works Clearinghouse (2020). *Standards handbook, version 4.1*. https://ies.ed.gov/ncee/wwc/Docs/referenceresources/WWC-Standards-Handbook-v4-1-508.pdf

Table 1

Implementation Metrics in Original and Replication Studies

Implementation Metric	Original Efficacy Study	Replication Study
Teacher Workshops	8 days (48 hours)	6 days (45 hours)
Coaching Visits per Classroom	1-2 per month	1-2 per month
Intervention Fidelity [a]	0.89	0.97
Classroom Curriculum Dosage [b]	Data not collected	85%
Home Curriculum Dosage [b]	81%	94%

[a] Maximum = 1.00; inadequate = .00 - .79; adequate = .80+; high = .90+

[b] Maximum = 100%; inadequate = 0%- 79%; adequate = 80%+; high = 90%+

Table 2

CMA and TEMA-3 Scores of Children in PK

			PK Fall		PK Spring	
Measure	Condition	*n*	M	(SD)	M	(SD)
CMA	Treatment	189	0.36	(0.17)	0.63	(0.16)
CMA	Control	189	0.36	(0.17)	0.54	(0.19)
TEMA-3	Treatment	182	5.43	(4.90)	14.24	(6.44)
TEMA-3	Control	183	5.18	(4.36)	12.19	(7.06)

Note. ITT sample

Table 3

Longitudinal Impacts on TEMA-3 in PK, K and 1

Wave	Impact estimate	Effect size	*SE*	*p*-value
2 (end of PK)	2.30**	0.36	0.89	0.011
3 (end of K)	2.23**	0.30	1.01	0.029
4 (end of 1)	1.44	0.17	1.12	0.199

Note: Impact estimate is calculated as the difference of least squared means. Effect size estimates differ from those reported in Supplementary Materials Table 2 due to differences in estimation methods.

** $p < .05$, two-tailed test.

Table 4

Baseline (Fall of PK) Equivalence on Pretest Measures

Outcome	Treatment N	Adjusted treatment mean	Unadjusted treatment *SD*	Control N	Adjusted control mean	Unadjusted control *SD*	Impact estimate	Pooled *SD*	Effect size (Cohen's d)	*p*-value
CMA	189	0.36	0.17	182	0.36	0.17	0.01	0.17	0.04	0.774
TEMA-3	189	5.43	4.90	183	5.18	4.36	0.25	4.64	0.05	0.669

Note: Treatment and control means were adjusted for clustering within PK classrooms.

Table 5

Impacts of Pre-K Mathematics on Math Achievement (Spring of PK)

Outcome	Treatment N	Adjusted treatment mean	Unadjusted treatment *SD*	Control N	Adjusted control mean	Unadjusted control *SD*	Impact estimate	Pooled *SD*	Effect size (Cohen's d)	*p*-value
CMA	189	0.63	0.16	182	0.54	0.19	0.09***	0.17	0.52	0.000
TEMA-3	189	14.24	6.44	183	12.19	7.06	2.05***	6.75	0.30	0.003

*** $p < .01$, two-tailed test.

Appendix A

Model Specification for Main Impact Analysis

(Equation 1) $$y_{ij} = \gamma_{00} + \gamma_{01} trt_{ij} + \sum \gamma_{n0} x_{nij} + r_{ij} + u_{0j},$$

where

y_{ij} is the outcome for student *i* in cluster *j* (CMA or TEMA-3);

γ_{00} is a constant term showing average student achievement in comparison clusters;

trt_j is a dummy variable indicating whether or not cluster *j* was randomly assigned to the treatment condition;

γ_{01} is the estimated effect of treatment;

x_{nij} is a vector of student level covariates: Fall prekindergarten pretest (CMA or TEMA-3), age, and gender);

γ_{n0} is a vector of coefficients associates with each of those covariates showing the association of each student-level characteristic and the outcome;

r_{ij} is an individual-level error term, assumed to have a normal distribution with mean zero and variance σ^2; and

$u_{0j}$is a cluster-level random error term, with an assumed normal distribution with mean zero and variance φ^2.

Effect Size Calculation Using Cohen's *d* Formula

(Equation 2) $$\frac{\widehat{mean}_T - \widehat{mean}_C}{SD_{pooled}},$$

where $\widehat{mean}_T$ equals the adjusted treatment group mean, $\widehat{mean}_C$ equals the adjusted control group mean, and SD_{pooled} is the pooled *SD*.

Formula Used to Calculate the Pooled *SD*

(Equation 3) $$SD_{pooled} = \sqrt{\frac{SD_T^2 + SD_C^2}{2}},$$

where SD_T equals the unadjusted treatment group SD and SD_C equals the unadjusted control group *SD*.