What Works Clearinghouse™

ies INSTITUTE OF EDUCATION SCIENCES

Early Childhood Education Updated December 2013

Pre-K Mathematics with DLM Early Childhood Express Math

Report Contents

Program Description[1]

Pre-K Mathematics is a supplemental curriculum designed to develop the informal mathematical knowledge and skills of preschool children with content organized into seven units. Specific math concepts and skills from each unit are taught in the classroom through teacher-guided, small-group activities with concrete manipulatives. The curriculum also includes take-home activities that parallel the small-group classroom activities, and are designed to help parents support their children's mathematical development at home. The *DLM Early Childhood Express Math* software includes corresponding math-based activities to reinforce math concepts taught in the classroom.

Research[2]

The What Works Clearinghouse (WWC) identified two studies of *Pre-K Mathematics* with *DLM Early Childhood Express Math* that both fall within the scope of the Early Childhood Education topic area and meet WWC evidence standards.[3] One study meets standards without reservations and one study meets WWC evidence standards with reservations. Together, they included 575 children in Head Start and public preschool programs in California and New York.

The WWC considers the extent of evidence for *Pre-K Mathematics* with *DLM Early Childhood Express Math* on the school readiness of preschool children to be small for three outcome domains—oral language, print knowledge, and phonological processing—and medium to large for one outcome domain—math. There were no studies that meet standards in two other domains, so we do not report on the effectiveness of *Pre-K Mathematics* with *DLM Early Childhood Express Math* for those domains in this intervention report. (See the Effectiveness Summary on p. 6 for further description of all domains.)

Effectiveness

Pre-K Mathematics with *DLM Early Childhood Express Math* was found to have no discernible effects on oral language, print knowledge, and phonological processing and positive effects on math for preschool children.

Table 1. Summary of findings[4]

Outcome domain	Rating of effectiveness	Improvement index *(percentile points)*		Number of studies	Number of students	Extent of evidence
		Average	Range			
Oral language	No discernible effects	+5	+3 to +7	1	296	Small
Print knowledge	No discernible effects	+3	0 to +7	1	297	Small
Phonological processing	No discernible effects	+2	na	1	270	Small
Math	Positive effects	+19	+6 to +32	2	575	Medium to large

na = not applicable

Program Information

Background

Developed by Alice Klein and Prentice Starkey with Alma Ramirez, *Pre-K Mathematics* is distributed by Pearson Scott Foresman, a division of Pearson Education as part of the Scott Foresman–Addison Wesley Mathematics (2008) series for grades prekindergarten through 6. Address: 1900 E. Lake Avenue, Glenview, IL 60025. Web: www.scottforesman.com. Telephone: (800) 552-2259.

The *DLM Early Childhood Express 2003, Math Resource Package* was developed by Douglas H. Clements and Julie Sarama, and includes a teacher's resource book (*DLM Early Childhood Express Math Resource Guide*) and the software. The resource guide and software are distributed as a set by McGraw-Hill School Education. Address: P.O. Box 182605, Columbus, OH 43218. Website: https://www.mheonline.com. Telephone: (800) 334-7344.

Program details

The *Pre-K Mathematics* curriculum includes activities organized in seven units containing closely-related math content: (a) counting and numbers, (b) understanding arithmetic operations (fall activities), (c) spatial sense and geometry, (d) patterns, (e) understanding arithmetic operations (spring activities), (f) measurement and data, and (g) logical reasoning. The program has both classroom activities with manipulatives and home activities with picture strips in English or Spanish. In the classroom, math concepts and skills are taught through teacher-guided, small-group activities over the course of the school year. Each week, a new math activity is presented twice to prekindergarten children in small-group lessons. Each group contains four to six children and the lessons last approximately 20 minutes. In addition, related math activities are sent home every 1 to 2 weeks and parents are encouraged to engage their children with these activities. Materials for the classroom and home are provided in a core program package that includes a teacher's curriculum book, Activity Aid Masters, Home Activity Masters, and Assessment Record Sheets.

The research evidence reviewed in this report addresses *Pre-K Mathematics* combined with *DLM Early Childhood Express Math* software. The software reinforces classroom support in two math domains: (a) geometric and spatial concepts and skills, and (b) numerical concepts and skills. It can be purchased only as part of a bundle called the *DLM Express Math Resource Package*, which includes the *DLM Early Childhood Express Math* software CD and the math resources guide booklet. The studies reviewed for this report used only the math software, and Klein et al. (2008) used only 27 of the math software activities included on the CD.

Cost

The *Pre-K Mathematics* Core Program Package (Pre-K Teacher's Edition, Activity Aid Masters, Home Activity Masters, and Assessment Record Sheets) costs $152.47, and the Spanish Home Activity Masters costs $16.97. Information on the cost of professional development is not available.

The *DLM Early Childhood Express 2003, Math Resource Package* with computer software costs $239.22.

Research Summary

The WWC identified eight studies that investigated the effects of *Pre-K Mathematics* with *DLM Early Childhood Express Math* on school readiness for preschool children.

Table 2. Scope of reviewed research

Grade	PK
Delivery method	Individual, Small group, Whole class
Program type	Curriculum

The WWC reviewed four of those studies against group design evidence standards. One study (Klein, Starkey, Clements, Sarama, & Iyer, 2008) is a randomized controlled trial that meets WWC evidence standards without reservations, and one study (Preschool Curriculum Evaluation Research [PCER] Consortium, 2008) is a randomized controlled trial that meets WWC evidence standards with reservations.[5] Those two studies are summarized in this report. Two studies do not meet WWC evidence standards. The remaining four studies do not meet WWC eligibility screens for review in this topic area. Citations for all eight studies are in the References section, which begins on p. 9.

Summary of study meeting WWC evidence standards without reservations

Klein et al. (2008) conducted a randomized controlled trial of 40 classrooms from six preschool programs in California and New York to assess the effectiveness of *Pre-K Mathematics* with *DLM Early Childhood Express Math*. Classrooms were recruited for the study as part of the pilot year of the national evaluation conducted by the PCER initiative.[6] The 40 classrooms were randomly assigned to either an intervention group implementing the *Pre-K Mathematics* curriculum with *DLM Early Childhood Express Math* as a supplement to the general curriculum already being used in the classroom, or a comparison group, which continued implementing the existing curriculum. During the 2002–03 school year (the pilot year of the PCER initiative), teachers in the intervention group implemented the *Pre-K Mathematics* curriculum with *DLM Early Childhood Express Math*. The study investigated the effects of the *Pre-K Mathematics* curriculum with *DLM Early Childhood Express Math* on the mathematics skills of 278 children (138 intervention and 140 comparison).[7]

Summary of study meeting WWC evidence standards with reservations

The PCER Consortium (2008) assessed the effects of *Pre-K Mathematics* with *DLM Early Childhood Express Math* using a randomized controlled trial of 40 Head Start and public preschool classrooms in California and New York. In the pilot year of the study (the 2002–03 school year), the classrooms were randomly assigned to either an intervention group, which implemented *Pre-K Mathematics* with *DLM Early Childhood Express Math* as a supplement to the general curriculum already being used in the classroom, or a comparison group, which used their existing curriculum. Thirty-four of the classrooms from the pilot year and six other classrooms were included in the sample for the national PCER evaluation year. One teacher was replaced.

Although the PCER Consortium (2008) study used a randomized controlled trial design to assign classrooms to intervention or comparison conditions in the pilot study year, the study analyzed data from the national PCER evaluation year (2003–04 school year), when children who had been in the classrooms at the time of random assignment (the start of the 2002–03 school year) had moved on to kindergarten and a new class of children had replaced them. In addition, seven teachers left the study and were replaced using processes that were not consistent with the original random assignment design. Thus, the study has high attrition at the child level and nonrandom replacement of some teachers, and under WWC standards, the study must demonstrate baseline equivalence between the intervention and comparison group sample of children used in the analyses of outcomes.

The authors investigated effects on oral language, print knowledge, phonological processing, and math. The outcome measures examined in this study were not examined in the Klein et al. (2008) study described above.

The WWC based its effectiveness ratings on findings from comparisons of 148 children who received *Pre-K Mathematics* with *DLM Early Childhood Express Math* and 149 comparison group children. Children in the sample were 4.3 years old (on average); 48% were male; 10% were reported as having a disability; 45% were African American, 23% were Hispanic, and 18% were Caucasian. The study demonstrated the baseline equivalence of the outcome measures in the oral language, print knowledge, phonological processing, and math domains for the analytic sample of intervention and comparison group children at the end of the preschool year.[8] The authors reported on the effects of *Pre-K Mathematics* with *DLM Early Childhood Express Math* in the spring of the preschool year and again at the end of kindergarten. The kindergarten findings are not reported here because information about the baseline equivalence of the outcome measures for the kindergarten sample was not provided in the report. The authors also reported findings on the Social Skills Rating Scale; however, these findings are not reported here because the current Early Childhood Education topic area protocol does not include sociobehavioral outcomes.

Effectiveness Summary

The WWC review of *Pre-K Mathematics* with *DLM Early Childhood Express Math* for the Early Childhood Education topic area includes child outcomes in six domains: oral language, print knowledge, phonological processing, early reading and writing, cognition, and math. The two studies of *Pre-K Mathematics* with *DLM Early Childhood Express Math* that meet WWC evidence standards reported findings in four of the six domains: (a) oral language, (b) print knowledge, (c) phonological processing, and (d) math. The findings below present the authors' estimates and WWC-calculated estimates of the size and statistical significance of the effects of *Pre-K Mathematics* with *DLM Early Childhood Express Math* on preschool children. For a more detailed description of the rating of effectiveness and extent of evidence criteria, see the WWC Rating Criteria on p. 19.

Summary of effectiveness for the oral language domain

One study that meets WWC standards with reservations reported findings in the oral language domain.

The PCER Consortium (2008) analyzed the effectiveness of *Pre-K Mathematics* with *DLM Early Childhood Express Math* on oral language using the Peabody Picture Vocabulary Test, Third Edition (PPVT-III) and the Test of Language Development–Primary: III (TOLD-P:3) Grammatic Understanding subtest. The authors reported that differences between the *Pre-K Mathematics* with *DLM Early Childhood Express Math* group and the comparison group were not statistically significant on any of these measures, and the WWC confirmed those findings. The average effect size across all findings was not large enough to be considered substantively important according to WWC criteria (that is, an effect size of at least 0.25). The WWC characterizes these study findings as an indeterminate effect.

Thus, for the oral language domain, one study showed an indeterminate effect. This results in a rating of no discernible effects, with a small extent of evidence.

Table 3. Rating of effectiveness and extent of evidence for the oral language domain

Rating of effectiveness	Criteria met
No discernible effects *No affirmative evidence of effects.*	In the one study that reported findings, the estimated impact of the intervention on outcomes in the *oral language* domain was neither statistically significant nor large enough to be substantively important.
Extent of evidence	**Criteria met**
Small	One study that included 296 children in 36 schools reported evidence of effectiveness in the *oral language* domain.

Summary of effectiveness for the print knowledge domain

One study that meets WWC standards with reservations reported findings in the print knowledge domain.

The PCER Consortium (2008) analyzed the effectiveness of *Pre-K Mathematics* with *DLM Early Childhood Express Math* on print knowledge using the Test of Early Reading Ability-III (TERA-III), the Woodcock-Johnson III (WJ-III) Letter-Word Identification subtest, and the WJ-III Spelling subtest. The authors reported that differences between the *Pre-K Mathematics* with *DLM Early Childhood Express Math* group and the comparison group were not statistically significant on any of these measures, and the WWC confirmed those calculations. The average effect size across all findings was not substantively important according to WWC criteria. The WWC characterizes these study findings as an indeterminate effect.

Thus, for the print knowledge domain, one study showed an indeterminate effect. This results in a rating of no discernible effects, with a small extent of evidence.

Table 4. Rating of effectiveness and extent of evidence for the print knowledge domain

Rating of effectiveness	Criteria met
No discernible effects *No affirmative evidence of effects.*	In the one study that reported findings, the estimated impact of the intervention on outcomes in the *print knowledge* domain was neither statistically significant nor large enough to be substantively important.
Extent of evidence	**Criteria met**
Small	One study that included 297 children in 36 schools reported evidence of effectiveness in the *print knowledge* domain.

Summary of effectiveness for the phonological processing domain

One study that meets WWC standards with reservations reported findings in the phonological processing domain.

The PCER Consortium (2008) analyzed the effectiveness of *Pre-K Mathematics* with *DLM Early Childhood Express Math* on phonological processing using the Preschool Comprehensive Test of Phonological and Print Processing (Pre-CTOPPP) Elision subtest. The authors reported that differences between the *Pre-K Mathematics* with *DLM Early Childhood Express Math* group and the comparison group were not statistically significant, and the WWC confirmed this calculation. The finding was not substantively important according to WWC criteria. The WWC characterizes this study finding as an indeterminate effect.

Thus, for the phonological processing domain, one study showed an indeterminate effect. This results in a rating of no discernible effects, with a small extent of evidence.

Table 5. Rating of effectiveness and extent of evidence for the phonological processing domain

Rating of effectiveness	Criteria met
No discernible effects *No affirmative evidence of effects.*	In the one study that reported findings, the estimated impact of the intervention on outcomes in the *phonological processing* domain was neither statistically significant nor large enough to be substantively important.
Extent of evidence	**Criteria met**
Small	One study that included 270 children in 36 schools reported evidence of effectiveness in the *phonological processing* domain.

Summary of effectiveness for the math domain

One study that meets WWC standards without reservations and one study that meets WWC standards with reservations reported findings in the math domain.

Klein et al. (2008) analyzed the effectiveness of *Pre-K Mathematics* with *DLM Early Childhood Express Math* on math using the Child Math Assessment (CMA). The authors reported that differences between the *Pre-K Mathematics* with *DLM Early Childhood Express Math* group and the comparison group were statistically significant. The WWC characterizes this study finding as a positive effect.

The PCER Consortium (2008) analyzed the effectiveness of *Pre-K Mathematics* with *DLM Early Childhood Express Math* on math using the WJ-III Applied Problems subtest, the Child Math Assessment–Abbreviated (CMA-A), and the Building Blocks Shape Composition task. The authors reported that differences between the *Pre-K Mathematics* with *DLM Early Childhood Express Math* group and the comparison group were statistically significant on the CMA-A and the Building Blocks Shape Composition task measures, and not statistically significant on the WJ-III Applied Problems measure. The WWC confirmed the statistical significance of the finding on the Building Blocks Shape Composition measure, but found the finding on the CMA-A to not be statistically significant after a correction

for multiple comparisons. The WWC characterizes the average effect size across all findings as a statistically significant positive effect.

Thus, for the math domain, two studies showed statistically significant positive effects. This results in a rating of positive effects, with a medium to large extent of evidence.

Table 6. Rating of effectiveness and extent of evidence for the math domain

Rating of effectiveness	Criteria met
Positive effects *Strong evidence of a positive effect with no overriding contrary evidence.*	In the two studies that reported findings, the estimated impact of the intervention on outcomes in the *math* domain was statistically significant and positive.
Extent of evidence	**Criteria met**
Medium to large	Two studies that included 575 children in 37 schools reported evidence of effectiveness in the *math* domain.

References

Study that meets WWC evidence standards without reservations

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.

Additional source:

Starkey, P., & Klein, A. (2005). *A longitudinal study of the effects of a pre-kindergarten mathematics curriculum on low-income children's mathematical knowledge*. Berkeley, CA: University of California.

Study that meets WWC evidence standards with reservations

Preschool Curriculum Evaluation Research (PCER) Consortium. (2008). Pre-K Mathematics supplemented with DLM Early Childhood Express Math software: University of California, Berkeley and University at Buffalo, State University of New York. In *Effects of preschool curriculum programs on school readiness* (pp. 131–142). Washington, DC: National Center for Education Research, Institute of Education Sciences, U.S. Department of Education.

Additional source:

Starkey, P., & Klein, A. (2005). *A longitudinal study of the effects of a pre-kindergarten mathematics curriculum on low-income children's mathematical knowledge*. Berkeley, CA: University of California.

Studies that do not meet WWC evidence standards

Clements, D. H., & Sarama, J. (2008). Experimental evaluation of the effects of a research-based preschool mathematics curriculum. *American Educational Research Journal, 45*(2), 443–494. The study does not meet WWC evidence standards because it uses a randomized controlled trial design that either did not generate groups using a random process or had nonrandom allocations after assignment, and the subsequent analytic intervention and comparison groups are not shown to be equivalent.

Starkey, P., Klein, A., & Wakeley, A. (2004). Enhancing young children's mathematical knowledge through a pre-kindergarten mathematics intervention. *Early Childhood Research Quarterly, 19*(1), 99–120. The study does not meet WWC evidence standards because because it uses a quasi-experimental design in which the analytic intervention and comparison groups are not shown to be equivalent.

Additional source:

Klein, A., & Starkey, P. (2004). Fostering preschool children's mathematical knowledge: Findings from the Berkeley Math Readiness Project. In D. H. Clements, J. Sarama, & A. M. DiBiase (Eds.), *Engaging young children in mathematics: Standards for early childhood mathematics education* (pp. 343–360). Mahwah, NJ: Lawrence Erlbaum Associates.

Studies that are ineligible for review using the Early Childhood Education Evidence Review Protocol

Malofeeva, E. V. (2008). Meta-analysis of mathematics instruction with young children. *Dissertation Abstracts International Section A: Humanities and Social Sciences, 69*(1-A), 113. The study is ineligible for review because it is a secondary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.

O'Dell, R. S. (2005). The nature of implementation of an innovative pre-K mathematics curriculum. *Dissertation Abstracts International Section A: Humanities and Social Sciences, 66*(08A), 249-2873. The study is ineligible for review because it does not include a student outcome.

Sarama, J., Clements, D., Starkey, P., Klein, A., & Wakeley, A. (2008). Scaling up the implementation of a pre-kindergarten mathematics curriculum: Teaching for understanding with trajectories and technologies (Pre-K Mathematics). *Journal of Research on Educational Effectiveness, 1*(2), 89–119. The study is ineligible for review because it does not examine the intervention implemented in a way that falls within the scope of the review—the intervention is bundled with other components.

Additional source:

Clements, D. H., & Sarama, J. (2006, June). *Scaling up the implementation of a pre-kindergarten mathematics curriculum: The Building Blocks curriculum*. Paper presented at the Institute of Education Sciences Research Conference, Washington, DC.

Starkey, P., & Klein, A. (2008). Sociocultural influences on young children's mathematical knowledge. In O. Saracho & B. Spodeck (Eds.), *Contemporary perspectives on mathematics in early childhood education* (pp. 253–276). Charlotte, NC: Information Age Publishing. The study is ineligible for review because it is a secondary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.

Appendix A.1: Research details for Klein et al. (2008)

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.

Table A1. Summary of findings

Meets WWC evidence standards without reservations

		Study findings	
Outcome domain	Sample size	Average improvement index *(percentile points)*	Statistically significant
Math	40 classrooms/278 children	+20	Yes

Setting

The study took place in 40 Head Start and state-funded preschool classrooms in California and New York. The classrooms were from four programs in California and two programs in New York.

Study sample

Twenty Head Start and 20 state-funded preschool classrooms were randomly assigned within program type either to an intervention group receiving *Pre-K Mathematics* with *DLM Early Childhood Express Math* or to a business-as-usual comparison group. Teachers volunteered to participate in the study, which took place during the 2002–03 school year. Of the 40 participating in the study, 33 were included in the PCER Consortium (2008) study that is described in Appendix A.2. The study began with 316 children from low-income families, combined across states. During the study, 38 children left, resulting in an analysis sample of 278 children. The mean age of the children in cohort 1 was 4.4 years. Fifty-three percent of the children were African American, 22% were Hispanic, 22% were Caucasian, 4% were Asian American, and 4% were interracial or another ethnicity. Forty-eight percent of the children were female.

Intervention group

Teachers implemented the *Pre-K Mathematics* curriculum classroom activities with small groups of four to six children in twice-a-week, 20-minute sessions for 29 weeks. Each week involved a new math activity. The small-group sessions included activities from the seven units of the curriculum: (a) counting and number, (b) understanding arithmetic operations (fall unit), (c) spatial sense and geometry, (d) patterns, (e) understanding arithmetic operations (spring unit), (f) measurement and data, and (g) logical reasoning. In addition, teachers supplemented the *Pre-K Mathematics* curriculum activities with two other instructional activities: (a) 27 computer activities based on the *DLM Early Childhood Express Math* software, and (b) mathematics learning centers, which included materials from the small-group activities and additional mathematics materials from the classroom. Home activity materials parallel to the classroom activities were sent home every 1 to 2 weeks for parents to use with their children. Teachers tracked children's progress using a Math Mastery Form, and intervention fidelity data were collected using the Fidelity of Implementation Record Sheet developed by the first two authors of the study.[9] In addition, teachers used a pre-existing general curriculum in their classrooms (curricula included *The Creative Curriculum*®, *High/Scope*, *Montessori*, or specialized literacy curricula and curricula developed by local teachers and school districts).

Comparison group

Teachers in the comparison group classrooms implemented the curriculum as business-as-usual in their programs. The curricula used in the comparison group classrooms included *The Creative Curriculum*®, *High/Scope*, *Montessori*, or specialized literacy curricula and curricula developed by local teachers and school districts.

Outcomes and measurement

The study measured intervention and comparison group children's mathematical knowledge with the researcher-developed CMA. All the children in the study were assessed at pretest before implementation of the intervention began in intervention classrooms, and at posttest after intervention classrooms completed implementation of the program. For a more detailed description of the CMA, see Appendix B.

Support for implementation

Teachers participated in a 4-day workshop at the beginning of the school year that focused on implementing the first three units of the curriculum and another 4-day workshop at mid-year that focused on implementing units four through seven. In addition, teachers were provided on-site training approximately twice a month and implementation fidelity checks once or twice each month.

Appendix A.2: Research details for PCER Consortium (2008)

Preschool Curriculum Evaluation Research (PCER) Consortium. (2008). Pre-K Mathematics supplemented with DLM Early Childhood Express Math software: University of California, Berkeley and University at Buffalo, State University of New York. In *Effects of preschool curriculum programs on school readiness* (pp. 131–142). Washington, DC: National Center for Education Research, Institute of Education Sciences, U.S. Department of Education.

Table A2. Summary of findings **Meets WWC evidence standards with reservations**

		Study findings	
Outcome domain	Sample size	Average improvement index *(percentile points)*	Statistically significant
Oral language	40 classrooms/296 children	+5	No
Print knowledge	40 classrooms/297 children	+3	No
Phonological processing	40 classrooms/270 children	+2	No
Math	40 classrooms/297 children	+18	Yes

Setting

The study took place in 40 Head Start and state-funded preschool classrooms in California and New York. The classrooms were from four programs in California and two programs in New York.

Study sample

The study, conducted during the 2003–04 and 2004–05 school years, included an intervention group implementing *Pre-K Mathematics* with *DLM Early Childhood Express Math* software and a comparison group. In the 2002–03 school year (the study's pilot year), teachers who volunteered to participate in the study were blocked by program type and randomly assigned to either an intervention group or a comparison group except for two pairs of classrooms, where randomization was conducted within the pair. In the study's evaluation year (2003–04), 33 of the 40 teachers were retained, and seven teachers were added using processes that were not consistent with the original random assignment design. In one site, three replacement classrooms were randomly selected, but the probability of selection to the intervention group was 66%, higher than the original 50% probability at the start of the study. In the other site, three teachers were nonrandomly placed into classrooms that were originally randomly assigned to intervention or comparison groups, and another teacher was randomly selected from among volunteers to fill an open teaching position in a study classroom. This resulted in a sample of 40 teachers

(20 intervention, 20 comparison) in the 2003–04 school year. Thus, for most of the classrooms, the intervention condition had been in place for a full year when the evaluation year started. After parental consent was obtained, the sample included 316 children at baseline; 297 children were included in the analytic sample (148 intervention, 149 comparison). Baseline equivalence between the analytic sample of intervention and comparison children was established based on baseline outcome measures data provided by the study authors. At baseline, children in the study classrooms averaged 4.3 years of age; 48% were male; 45% were African American, 23% were Hispanic, and 18% were Caucasian. Ten percent of the children were identified as having a disability.

Intervention group

Teachers conducted mathematics activities from *Pre-K Mathematics* twice a week with groups of four to six children for approximately 20 minutes per group. During each classroom session, teachers completed Assessment Record Sheets that were tied to the mathematics activity in that session. Twenty-nine classroom activities were completed, and teachers sent 19 home activities and materials for children to complete at home. Materials for home mathematics activities were sent home every 1 to 2 weeks. *Pre-K Mathematics* was supplemented with the *DLM Early Childhood Express Math* software, which included 26 numerical, quantitative, geometric, and spatial activities. The software program provided individualized mathematics instructional activities approximately twice a week. The intervention took place over 36 weeks. In addition, teachers used a pre-existing general curriculum in their classrooms (curricula included *The Creative Curriculum®*, *High/Scope*, *Montessori*, or specialized literacy curricula and curricula developed by local teachers and school districts).

Comparison group

The business-as-usual comparison group participated in the curriculum used in their programs, such as *The Creative Curriculum®*, *High/Scope*, *Montessori*, or specialized literacy curricula and curricula developed by local teachers and school districts.

Outcomes and measurement

The outcome domains of oral language, print knowledge, phonological processing, and math were assessed with standardized measures. Oral language was assessed with the Peabody Picture Vocabulary Test, Third Edition (PPVT-III) and the Grammatic Understanding subtest of the TOLD-P:3. Print knowledge was assessed with the TERA-III and the WJ-III Letter-Word Identification and Spelling subtests. Phonological processing was assessed with the Pre-CTOPPP Elision subtest. Math was assessed with the WJ-III Applied Problems subtest, the CMA-A, and the Building Blocks Shape Composition test. The pretest assessment was conducted in the fall of the 2003–04 school year, and the posttest assessment in the spring of the same school year, when the children in the sample attended preschool. Trained research staff administered all assessments. For a more detailed description of these outcome measures, see Appendix B.

Support for implementation

During the 2002–03 school year (the evaluation year), intervention teachers participated in a 4-day training workshop. Ongoing on-site training was provided approximately twice per month. Prior to the 2003–04 school year, teachers in the intervention group received a 2-day refresher workshop. Project staff observed and rated implementation fidelity of the small group sessions once or twice a month and provided feedback to teachers.

Appendix B: Outcome measures for each domain

Oral language	
Peabody Picture Vocabulary Test, Third Edition (PPVT-III)	A nationally-standardized, individually-administered assessment of children's receptive vocabulary in which children demonstrate understanding of a spoken word by pointing to a picture that best represents the meaning (as cited in PCER Consortium, 2008).
Test of Language Development–Primary: III (TOLD-P:3) Grammatic Understanding subtest	A nationally-standardized, individually-administered assessment of children's ability to comprehend the meaning of sentences by selecting pictures that most accurately represent the sentence (as cited in PCER Consortium, 2008).
Print knowledge	
Test of Early Reading Ability–III (TERA-III)	A nationally-standardized, individually-administered assessment of children's developing reading skills with three subtests: alphabet, conventions, and meaning (as cited in PCER Consortium, 2008).
Woodcock-Johnson III (WJ-III) Letter-Word Identification subtest	A nationally-standardized, individually-administered assessment of children's identification of letters and reading of words (as cited in PCER Consortium, 2008).
WJ-III Spelling subtest	A nationally-standardized, individually-administered assessment of children's prewriting skills, such as drawing lines, tracing, and writing letters (as cited in PCER Consortium, 2008).
Phonological processing	
Preschool Comprehensive Test of Phonological and Print Processing (Pre-CTOPPP), Elision subtest	An individually-administered assessment of children's ability to identify and manipulate sounds in spoken words, using word prompts and picture plates for the first nine items and word prompts only for later items (as cited in PCER Consortium, 2008).
Math	
Building Blocks Shape Composition task	An individually-administered assessment of early mathematics achievement, this measure was modified for PCER from the Early Maths Assessment, developed by Clements, Sarama, and Liu (2008).[10] Children use blocks to fill in a puzzle and are assessed on whether they fill the puzzle without gaps or hangovers (as cited in PCER Consortium, 2008).
Child Math Assessment (CMA)	A researcher-developed measure designed to assess young children's early mathematical knowledge in the areas of number, arithmetic, space and geometry, measurement, and pattern knowledge (as cited in PCER Consortium, 2008).[11]
Child Math Assessment–Abbreviated (CMA-A) Composite Score	An individually-administered assessment of early mathematics achievement, this measure is the average of four subscales: (a) solving addition and subtraction problems using visible objects, (b) constructing a set of objects equal in number to a given set, (c) recognizing shapes, and (d) copying a pattern using objects that vary in color and identity from the model pattern. This assessment was adapted for PCER from a more comprehensive early mathematics assessment by Klein and Starkey (2002), who also developed the *Pre-K Mathematics* curriculum and participated in one of the research teams for PCER (as cited in PCER Consortium, 2008).[12]
Research-based Early Mathematics Assessment	A researcher-developed measure that uses individual child interviews to assess young children's mathematical knowledge and skills in the areas of number, geometry, measurement, and patterning (as cited in Sarama et al., 2008).
WJ-III Applied Problems subtest	A nationally-standardized, individually-administered assessment of children's ability to solve numerical and spatial problems, presented verbally with accompanying pictures of objects (as cited in PCER Consortium, 2008).

Appendix C.1: Findings included in the rating for the oral language domain

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			*p*-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
PCER Consortium, 2008[a]								
PPVT-III	Preschool children	40 classrooms/ 296 children	nr	nr	nr	0.18	+7	> 0.05
TOLD-P:3 Grammatic Understanding subtest	Preschool children	40 classrooms/ 269 children	nr	nr	nr	0.07	+3	> 0.05
Domain average for oral language (PCER Consortium, 2008)						**0.13**	**+5**	**Not statistically significant**
Domain average for oral language across all studies						**0.13**	**+5**	**na**

Table Notes: For mean difference, effect size, and improvement index values reported in the table, a positive number favors the intervention group and a negative number favors the comparison group. The effect size is a standardized measure of the effect of an intervention on child outcomes, representing the average change expected for all children who are given the intervention (measured in standard deviations of the outcome measure). The improvement index is an alternate presentation of the effect size, reflecting the change in an average child's percentile rank that can be expected if the child is given the intervention. The WWC-computed average effect size is a simple average rounded to two decimal places; the average improvement index is calculated from the average effect size. The statistical significance of the study's domain average was determined by the WWC. nr = not reported. na = not applicable. PPVT-III = Peabody Picture Vocabulary Test, Third Edition. TOLD-P:3 = Test of Language Development–Primary: III.

[a] For PCER Consortium (2008), a correction for multiple comparisons was needed but did not affect whether any of the contrasts were found to be statistically significant. The effect sizes and *p*-values presented here were reported in the original study (in Table A-18, based on an alternative estimation approach, ANCOVA, that included the baseline pretest). Mean scores and differences are not reported in this table because the study-reported group means and differences were not adjusted for baseline pretest scores. This study is characterized as having an indeterminate effect because the mean effect is neither statistically significant nor substantively important, accounting for multiple comparisons.

Appendix C.2: Findings included in the rating for the print knowledge domain

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			*p*-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
PCER Consortium, 2008[a]								
TERA-III	Preschool children	40 classrooms/ 270 children	nr	nr	nr	0.00	0	> 0.05
WJ-III Letter-Word Identification subtest	Preschool children	40 classrooms/ 297 children	nr	nr	nr	0.06	+2	> 0.05
WJ-III Spelling subtest	Preschool children	40 classrooms/ 297 children	nr	nr	nr	0.17	+7	> 0.05
Domain average for print knowledge (PCER Consortium, 2008)						**0.08**	**+3**	**Not statistically significant**
Domain average for print knowledge across all studies						**0.08**	**+3**	**na**

Table Notes: For mean difference, effect size, and improvement index values reported in the table, a positive number favors the intervention group and a negative number favors the comparison group. The effect size is a standardized measure of the effect of an intervention on child outcomes, representing the average change expected for all children who are given the intervention (measured in standard deviations of the outcome measure). The improvement index is an alternate presentation of the effect size, reflecting the change in an average child's percentile rank that can be expected if the child is given the intervention. The WWC-computed average effect size is a simple average rounded to two decimal places; the average improvement index is calculated from the average effect size. The statistical significance of the study's domain average was determined by the WWC. nr = not reported. na = not applicable. TERA-III = Test of Early Reading Ability–III. WJ-III = Woodcock-Johnson III.

[a] For PCER Consortium (2008), a correction for multiple comparisons was needed but did not affect whether any of the contrasts were found to be statistically significant. The effect sizes and *p*-values presented here were reported in the original study (in Table A-18, based on an alternative estimation approach, ANCOVA, that included the baseline pretest). Mean scores and differences are not reported in this table because the study-reported group means and differences were not adjusted for baseline pretest scores. This study is characterized as having an indeterminate effect because the mean effect is neither statistically significant nor substantively important, accounting for multiple comparisons.

Appendix C.3: Findings included in the rating for the phonological processing domain

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			*p*-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
PCER Consortium, 2008[a]								
Pre-CTOPPP Ellision subtest	Preschool children	40 classrooms/ 270 children	9.67 (3.87)	9.24 (4.35)	0.43	0.04	+2	> 0.05
Domain average for phonological processing (PCER Consortium, 2008)						**0.04**	**+2**	**Not statistically significant**
Domain average for phonological processing across all studies						**0.04**	**+2**	**na**

Table Notes: For mean difference, effect size, and improvement index values reported in the table, a positive number favors the intervention group and a negative number favors the comparison group. The effect size is a standardized measure of the effect of an intervention on child outcomes, representing the average change expected for all children who are given the intervention (measured in standard deviations of the outcome measure). The improvement index is an alternate presentation of the effect size, reflecting the change in an average child's percentile rank that can be expected if the child is given the intervention. na = not applicable. Pre-CTOPPP = Preschool Comprehensive Test of Phonological and Print Processing.

[a] For PCER Consortium (2008), the effect size, mean difference, and *p*-value presented here was reported in the original study (in Table 10.4, Table D-12a, and Table 10.4, respectively). Adjustment for the baseline pretest scores was not required for this domain. Thus, the intervention group mean equals the sum of the unadjusted comparison group mean and the covariate-adjusted mean difference reported in the study (in Table C-12a and Table D-12a, respectively). This study is characterized as having an indeterminate effect because the mean effect is neither statistically significant nor substantively important.

Appendix C.4: Findings included in the rating for the math domain

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			*p*-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
Klein et al., 2008[a]								
CMA	Preschool children	40 classrooms/ 278 children	0.55 (0.13)	0.47 (0.14)	0.07	0.52	+20	0.00
Domain average for math (Klein et al., 2008)						**0.52**	**+20**	**Statistically significant**
PCER Consortium, 2008[b]								
WJ-III Applied Problems subtest	Preschool children	40 classrooms/ 296 children	nr	nr	nr	0.16	+6	> 0.05
CMA-A Composite	Preschool children	40 classrooms/ 297 children	nr	nr	nr	0.35	+14	0.05
Building Blocks Shape Composition task	Preschool children	40 classrooms/ 297 children	nr	nr	nr	0.91	+32	> 0.01
Domain average for math (PCER Consortium, 2008)						**0.47**	**+18**	**Statistically significant**
Domain average for math across all studies						**0.50**	**+19**	**na**

Table Notes: For mean difference, effect size, and improvement index values reported in the table, a positive number favors the intervention group and a negative number favors the comparison group. The effect size is a standardized measure of the effect of an intervention on child outcomes, representing the average change expected for all children who are given the intervention (measured in standard deviations of the outcome measure). The improvement index is an alternate presentation of the effect size, reflecting the change in an average child's percentile rank that can be expected if the child is given the intervention. The WWC-computed average effect size is a simple average rounded to two decimal places; the average improvement index is calculated from the average effect size. The statistical significance of each study's domain average was determined by the WWC. nr = not reported. na = not applicable. CMA = Child Math Assessment. WJ-III = Woodcock-Johnson III. CMA-A = Child Math Assessment–Abbreviated.

[a] For Klein et al. (2008), no corrections for clustering or multiple comparisons were needed. The mean difference, effect size, and *p*-value presented here were reported in the original study. This study is characterized as having a statistically significant positive effect because the effect for at least one measure within the domain is positive and statistically significant, and no effects are negative and statistically significant, accounting for multiple comparisons.

[b] For PCER Consortium (2008), a correction for multiple comparisons was needed and resulted in significance levels that differ from those in the original study. The effect sizes and *p*-values presented here were reported in the original study (in Table A-18, based on an alternative estimation approach, ANCOVA, that included the baseline pretest). Due to the multiple comparisons correction, the *p*-value of 0.05 for the *CMA-A Composite* outcome measure was higher than the critical *p*-value for statistical significance; therefore, the WWC does not find the result to be statistically significant. Mean scores and differences are not reported in this table because the study-reported group means and differences were not adjusted for baseline pretest scores. This study is characterized as having a statistically significant positive effect because the effect for at least one measure within the domain is positive and statistically significant, and no effects are negative and statistically significant, accounting for multiple comparisons.

Endnotes

[1] The descriptive information for this program was obtained from publicly available sources: the program's website: (http://pearsonschool.com, downloaded February 2012) and the research literature (Klein et al., 2008; PCER Consortium, 2008; and Sarama et al., 2008). The WWC requests developers review the program description sections for accuracy from their perspective. The program description was provided to the developer in March 2012; however, the WWC received no response. Further verification of the accuracy of the descriptive information for this program is beyond the scope of this review. The literature search reflects documents publicly available by December 2012.

[2] The previous report was released in July 2007. This report has been updated to include reviews of six studies that have been released since 2007. Of the additional studies, three were not within the scope of the review protocol for the Early Childhood Education topic area. A complete list and disposition of all studies reviewed are provided in the references. The studies in this report were reviewed using the Evidence Standards from the WWC Procedures and Standards Handbook (version 2.1), along with those described in the Early Childhood Education review protocol (version 2.0). The evidence presented in this report is based on available research. Findings and conclusions may change as new research becomes available.

[3] Absence of conflict of interest: The PCER Consortium (2008) study summarized in this intervention report was prepared by staff of one of the WWC contractors. Because the principal investigator for the WWC review of early childhood education is also a staff member of that contractor, the study was rated by staff members from a different organization, who also prepared the intervention report. The report was then reviewed by the principal investigator, a WWC Quality Assurance reviewer, and an external peer reviewer.

[4] For criteria used in the determination of the rating of effectiveness and extent of evidence, see the WWC Rating Criteria on p. 19. These improvement index numbers show the average and range of child-level improvement indices for all findings across the studies.

[5] These two studies examine the effects of *Pre-K Mathematics* with *DLM Early Childhood Express Math* based on data from the PCER initiative. However, each study examines a unique sample of preschool children and examines different outcome measures. Klein et al. (2008) examines the effects of that intervention for a sample of children attending preschool classrooms during the pilot year (first year) of the PCER initiative. PCER Consortium (2008) examines the effects of that intervention for a sample of children attending preschool during the second year of the PCER initiative (the 2003–04 school year). Thirty-three teachers participated in both the pilot and the second year of the PCER initiative, and seven teachers who were included in the first-year study left the study and were replaced for its second year. The average age of the children in the sample examined by Klein et al. (2008) and PCER Consortium (2008) is 4.4 and 4.3 years, respectively.

[6] The national PCER Consortium (2008) study conducted a rigorous efficacy evaluation of 14 preschool curricula. Twelve research teams implemented one or two curricula in preschool settings serving predominantly low-income children using an experimental design. For each team, preschools or classrooms were randomly assigned to either the intervention curricula or the comparison curricula, and the children were followed from prekindergarten through kindergarten. The studies each used a common set of measures, with the cohort of children beginning preschool in the summer/fall of 2003. The PCER Consortium (2008) summarized the details and results of each curriculum study.

[7] This information was provided by the study authors, at the WWC's request.

[8] An author query was conducted to obtain the study data necessary to establish equivalence at baseline for one outcome measure in each domain (i.e., unadjusted means and standard deviations of the outcome measures for the intervention and the comparison groups). The pretest data provided for each domain were used to establish baseline equivalence for the domain.

[9] Klein, A., & Starkey, P. (2002). *Pre-K Mathematics curriculum*. Glenview, IL: Scott Foresman.

[10] The CMA was developed by the researchers, who are also the program developers, and this measure was developed for the purposes of this research project. The measure was confirmed to have sufficient face validity by the WWC Early Childhood Education principal investigator and by a psychometric study to establish its measurement properties by the study authors as a part of the Institute of Education Sciences-funded Interagency Education Research Initiative (IERI) Scale-Up project.

[11] Klein, A., & Starkey, P. (2002). *Child Math Assessment-Abbreviated*. Berkeley, CA: Author.

[12] Clements, D. H., Sarama, J., & Liu, X. (2008). Development of a measure of early mathematics achievement using the Rasch model: The Research-based Early Maths Assessment. *Educational Psychology, 28*(4), 457–482.

Recommended Citation

U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse. (2013, December). *Early Childhood Education intervention report: Pre-K Mathematics with DLM Early Childhood Express Math*. Retrieved from http://whatworks.ed.gov

WWC Rating Criteria

Criteria used to determine the rating of a study

Study rating	Criteria
Meets WWC evidence standards without reservations	A study that provides strong evidence for an intervention's effectiveness, such as a well-implemented RCT.
Meets WWC evidence standards with reservations	A study that provides weaker evidence for an intervention's effectiveness, such as a QED or an RCT with high attrition that has established equivalence of the analytic samples.

Criteria used to determine the rating of effectiveness for an intervention

Rating of effectiveness	Criteria
Positive effects	Two or more studies show statistically significant positive effects, at least one of which met WWC evidence standards for a strong design, AND No studies show statistically significant or substantively important negative effects.
Potentially positive effects	At least one study shows a statistically significant or substantively important positive effect, AND No studies show a statistically significant or substantively important negative effect AND fewer or the same number of studies show indeterminate effects than show statistically significant or substantively important positive effects.
Mixed effects	At least one study shows a statistically significant or substantively important positive effect AND at least one study shows a statistically significant or substantively important negative effect, but no more such studies than the number showing a statistically significant or substantively important positive effect, OR At least one study shows a statistically significant or substantively important effect AND more studies show an indeterminate effect than show a statistically significant or substantively important effect.
Potentially negative effects	One study shows a statistically significant or substantively important negative effect and no studies show a statistically significant or substantively important positive effect, OR Two or more studies show statistically significant or substantively important negative effects, at least one study shows a statistically significant or substantively important positive effect, and more studies show statistically significant or substantively important negative effects than show statistically significant or substantively important positive effects.
Negative effects	Two or more studies show statistically significant negative effects, at least one of which met WWC evidence standards for a strong design, AND No studies show statistically significant or substantively important positive effects.
No discernible effects	None of the studies shows a statistically significant or substantively important effect, either positive or negative.

Criteria used to determine the extent of evidence for an intervention

Extent of evidence	Criteria
Medium to large	The domain includes more than one study, AND The domain includes more than one school, AND The domain findings are based on a total sample size of at least 350 students, OR, assuming 25 students in a class, a total of at least 14 classrooms across studies.
Small	The domain includes only one study, OR The domain includes only one school, OR The domain findings are based on a total sample size of fewer than 350 students, AND, assuming 25 students in a class, a total of fewer than 14 classrooms across studies.

Glossary of Terms

Attrition Attrition occurs when an outcome variable is not available for all participants initially assigned to the intervention and comparison groups. The WWC considers the total attrition rate and the difference in attrition rates across groups within a study.

Clustering adjustment If intervention assignment is made at a cluster level and the analysis is conducted at the student level, the WWC will adjust the statistical significance to account for this mismatch, if necessary.

Confounding factor A confounding factor is a component of a study that is completely aligned with one of the study conditions, making it impossible to separate how much of the observed effect was due to the intervention and how much was due to the factor.

Design The design of a study is the method by which intervention and comparison groups were assigned.

Domain A domain is a group of closely related outcomes.

Effect size The effect size is a measure of the magnitude of an effect. The WWC uses a standardized measure to facilitate comparisons across studies and outcomes.

Eligibility A study is eligible for review and inclusion in this report if it falls within the scope of the review protocol and uses either an experimental or matched comparison group design.

Equivalence A demonstration that the analysis sample groups are similar on observed characteristics defined in the review area protocol.

Extent of evidence An indication of how much evidence supports the findings. The criteria for the extent of evidence levels are given in the WWC Rating Criteria on p. 19.

Improvement index Along a percentile distribution of students, the improvement index represents the gain or loss of the average student due to the intervention. As the average student starts at the 50th percentile, the measure ranges from –50 to +50.

Multiple comparison adjustment When a study includes multiple outcomes or comparison groups, the WWC will adjust the statistical significance to account for the multiple comparisons, if necessary.

Quasi-experimental design (QED) A quasi-experimental design (QED) is a research design in which subjects are assigned to intervention and comparison groups through a process that is not random.

Randomized controlled trial (RCT) A randomized controlled trial (RCT) is an experiment in which investigators randomly assign eligible participants into intervention and comparison groups.

Rating of effectiveness The WWC rates the effects of an intervention in each domain based on the quality of the research design and the magnitude, statistical significance, and consistency in findings. The criteria for the ratings of effectiveness are given in the WWC Rating Criteria on p. 19.

Single-case design A research approach in which an outcome variable is measured repeatedly within and across different conditions that are defined by the presence or absence of an intervention.

Standard deviation The standard deviation of a measure shows how much variation exists across observations in the sample. A low standard deviation indicates that the observations in the sample tend to be very close to the mean; a high standard deviation indicates that the observations in the sample tend to be spread out over a large range of values.

Statistical significance Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups. The WWC labels a finding statistically significant if the likelihood that the difference is due to chance is less than 5% ($p < 0.05$).

Substantively important A substantively important finding is one that has an effect size of 0.25 or greater, regardless of statistical significance.

Please see the WWC Procedures and Standards Handbook (version 2.1) for additional details.