

# Appendix

## Appendix A1.1 Study characteristics: Clements & Sarama, 2006 (randomized controlled trial)

Characteristic	Description
<b>Study citation</b>	Clements, D. H., & Sarama, J. (2006, June). <i>Scaling up the implementation of a pre-Kindergarten mathematics curriculum: The Building Blocks curriculum</i> . Paper presented at the Institute of Education Sciences Research Conference, Washington, D.C.
<b>Participants</b>	Teachers were randomly assigned to conditions in two separate steps. Twenty-four teachers from preschool programs serving low-income children were randomly assigned to two intervention groups ( <i>Building Blocks for Math</i> or <i>Pre-K Mathematics</i> ) <sup>1</sup> or a business-as-usual comparison group. Twelve teachers from preschool programs serving mixed-income children were randomly assigned to the <i>Building Blocks for Math</i> group or the business-as-usual comparison group. Consequently there were a total of 14 teachers in the <i>Building Blocks for Math</i> group and 14 teachers in the business-as-usual comparison group. Eight preschool-age children were randomly selected from each classroom for assessment (N = 224). <sup>2</sup> After attrition, the final sample included 28 teachers and 202 children (14 teachers and 101 children in the <i>Building Blocks for Math</i> group; 14 teachers and 101 children in the business-as-usual comparison group).
<b>Setting</b>	The study was conducted in Head Start and state-funded preschool programs in New York State.
<b>Intervention</b>	Children in the <i>Building Blocks for Math</i> intervention condition used the <i>Building Blocks for Math</i> curriculum in 10- to 15-minute small-group (4–6 children) math activities weekly. These children also participated in 5- to 15-minute whole-group math activities four times a week and 5- to 10-minute computer activities ( <i>DLM Express</i> ) twice a week. Related family activities were sent home weekly. The intervention lasted for 26 weeks, and intervention teachers maintained their daily activities and schedule while inserting mathematics activities at appropriate times during the day. <sup>3</sup>
<b>Comparison</b>	Children in the business-as-usual comparison group participated in their regular daily activities and schedule, with emphasis on small groups and computer activities. These included city-wide math activities, <i>Creative Curriculum</i> , Montessori math activities, or “home-grown” math materials based on state standards.
<b>Primary outcomes and measurement</b>	The primary outcome domain assessed was math and it was measured with the Early Mathematics Assessment (see Appendix A2 for a more detailed description of the outcome measure). The study authors also assessed implementation fidelity with the Fidelity of Implementation measure and the quality of the mathematics environment using the Classroom Observation of Early Mathematics Environment and Teaching. This WWC review does not include the results from these observations in this WWC review. <sup>4</sup>
<b>Teacher training</b>	Professional development activities for teachers in the <i>Building Blocks for Math</i> group consisted of four days of training, a monthly two-hour class, and monthly in-class coaching by project staff. Teacher training covered a number of topics such as supporting mathematical development in the classroom, recognizing and supporting mathematics throughout the day, setting up mathematics learning centers, teaching with computers, small-group activities, and supporting mathematical development in the home. Learning trajectories were emphasized in the <i>Building Blocks for Math</i> training via a web-based application called <i>Building Blocks Learning Trajectories</i> .

1. The study also included a *Pre-K Mathematics* intervention group, which used *DLM Express* as an additional component. The study authors labeled the *Pre-K Mathematics* group as the “comparison group” and the *Building Blocks for Math* group as the “intervention group”; however, the WWC considers *Pre-K Mathematics* as a separate intervention (see the separate [WWC Pre-K Mathematics intervention report](#)). For the rating of effectiveness in this WWC intervention report, the WWC includes only the results comparing the *Building Blocks for Math* group to the business-as-usual comparison group; however, results for the comparison between the curricula are included in Appendix A4.
2. The remaining eight teachers were assigned to the *Pre-K Mathematics* group, which is not the main focus of this WWC intervention report.
3. Children in the *Pre-K Mathematics* intervention group participated in 15- to 20-minute small-group activities at least twice a week and 10- to 15-minute whole-class math activities once a week. In addition, these classrooms used *DLM Express* software for 5 to 10 minutes twice a week. Weekly letters were sent to parents that included math activities similar to those children were learning at school. The intervention lasted for 26 weeks and the teachers maintained their daily activities and schedule while inserting mathematics activities at appropriate times during the day.
4. For further details about the outcomes included in the Early Childhood Education topic review, please see the [Early Childhood Education Protocol](#).

## Appendix A1.2 Study characteristics: Clements & Sarama, 2007 (randomized controlled trial)<sup>1</sup>

Characteristic	Description
<b>Study citation</b>	<p>Clements, D. H., &amp; Sarama, J. (2007). Effects of a preschool mathematics curriculum: Summative research on the Building Blocks project. <i>Journal for Research in Mathematics Education</i>, 38(2), 136–163.</p> <p><i>Additional sources:</i></p> <p>Clements, D. H., &amp; Sarama, J. (2002). <i>Effects of a preschool mathematics curriculum: Research on the NSF-funded Building Blocks Project</i>. University at Buffalo, State University of New York.</p> <p>Clements, D. H., &amp; Sarama, J. (n.d.). <i>Effects of a preschool mathematics curriculum: Summary research on the NSF-funded Building Blocks project</i>. University at Buffalo, State University of New York. Retrieved from <a href="http://www.gse.buffalo.edu/org/buildingblocks/writings/Building%20Blocks%20Research%201.pdf">http://www.gse.buffalo.edu/org/buildingblocks/writings/Building%20Blocks%20Research%201.pdf</a></p>
<b>Participants</b>	Two teachers from a Head Start center and two teachers from a state-funded preschool program were randomly assigned to the <i>Building Blocks for Math</i> group or the business-as-usual comparison group. There were 77 children in the four classrooms; however, nine children left the classrooms during the course of the study. Final analysis samples ranged from 53 to 61 children. <sup>1</sup> The majority of the children were from low-income families. The mean age of the children was 49.9 months, and 49% of the children were female.
<b>Setting</b>	The study took place in one Head Start and one state-funded preschool program in New York State.
<b>Intervention</b>	The intervention group used <i>Building Blocks for Math</i> for 25 weeks. <i>Building Blocks for Math</i> includes whole- and small-group activities and games, free-choice learning centers, ideas for integrating mathematics throughout the school day, computer software ( <i>DLM Express</i> ), and books, game sheets, and manipulatives. Teachers in the state-funded preschool program implemented the curriculum with a high degree of fidelity, whereas teachers in the Head Start center implemented it with a moderate—but adequate—degree of fidelity.
<b>Comparison</b>	Children in the business-as-usual comparison group used their typical preschool math curricula and activities over the same 25 weeks. The typical preschool math curriculum used in the Head Start comparison classroom was <i>Creative Curriculum</i> and math activities developed by the program. The state-funded preschool program comparison classroom continued implementing the school's regular math curriculum.
<b>Primary outcomes and measurement</b>	The primary outcome domain assessed was math, and it was measured with the Number and Geometry sections of the Building Blocks Assessment of Early Mathematics, PreK–K (see Appendix A2 for more detailed descriptions of outcome measures).
<b>Teacher training<sup>2</sup></b>	Intervention teachers participated with researchers in a half-day review of curriculum materials. Project staff observed intervention teachers implementing the curriculum and were available to answer questions and discuss implementation issues with teachers as needed.

1. According to data provided by the study authors upon the WWC request, there was within-cluster attrition of 21.6% between the intervention (10% of children) and business-as-usual comparison (31.6% of children) groups. However, the WWC did not downgrade the study because there was pretest equivalence between groups for the non-attriters.
2. The study authors had worked previously with the teachers at the state-funded preschool program, the intervention teacher at the Head Start was inexperienced, and the comparison teacher at Head Start had taught Head Start for eight years.

## Appendix A2 Outcome measures in the math domain<sup>1</sup>

Outcome measure	Description
<b>Early Mathematics Assessment</b>	A researcher-developed measure that uses two individual child interviews to assess young children's mathematical knowledge and skills in the areas of number, geometry, measurement, and patterning (as cited in Clements & Sarama, 2006).
<b>Building Blocks Assessment of Early Mathematics, PreK–K: Number section</b>	A researcher-developed measure that uses individual child interviews to assess young children's mathematical knowledge. The number section measures nine topics, including verbal counting, object counting, number recognition, number comparison, number sequencing, numerals, number composition, adding and subtracting, and place value (as cited in Clements & Sarama, 2007).
<b>Building Blocks Assessment of Early Mathematics, PreK–K: Geometry section</b>	A researcher-developed measure that uses individual child interviews to assess young children's mathematical knowledge. The geometry section measures seven topics, including shape identification, shape composition, congruence, construction of shape, turns, measurement, and patterning (as cited in Clements & Sarama, 2007).

1. The Early Mathematics Assessment is an earlier version of the Building Blocks Assessment of Early Mathematics. The name change was at the publisher's request.

## Appendix A3 Summary of study findings included in the rating for the math domain<sup>1</sup>

Outcome measure	Study sample	Sample size (teachers/ children) <sup>3</sup>	Authors' findings from the study			WWC calculations		
			Mean outcome (standard deviation <sup>2</sup> )		Mean difference <sup>5</sup> ( <i>Building Blocks for Math</i> – comparison)	Effect size <sup>6</sup>	Statistical significance <sup>7</sup> (at $\alpha = 0.05$ )	Improvement index <sup>8</sup>
			<i>Building Blocks for Math</i> group <sup>4</sup>	Comparison group <sup>4</sup>				
<b>Clements &amp; Sarama, 2006 (randomized controlled trial)<sup>9</sup></b>								
Early Mathematics Assessment	Preschool children	28/202	61.69 (7.46)	53.22 (8.38)	8.47	1.06	Statistically significant	+36
<b>Average<sup>10</sup> for math (Clements &amp; Sarama, 2006)</b>						1.06	Statistically significant	+36
<b>Clements &amp; Sarama, 2007 (randomized controlled trial)<sup>11</sup></b>								
BBA: Number section	4 year olds	4/61	28.23 (14.47)	17.93 (12.48)	10.30	0.75	ns	+27
BBA: Geometry section	4 year olds	4/53	15.75 (3.81)	10.64 (3.35)	5.11	1.40	Statistically significant	+42
<b>Average<sup>10</sup> for math (Clements &amp; Sarama, 2007)</b>						1.08	Statistically significant	+36
<b>Domain average<sup>10</sup> for math across all studies</b>						1.07	na	+36

BBA = Building Blocks Assessment of Early Mathematics, PreK–K

ns = not statistically significant

na = not applicable

1. This appendix reports findings considered for the effectiveness rating and the average improvement indices. Findings from the head-to-head comparison of *Building Blocks for Math* to *Pre-K Mathematics* from Clements and Sarama (2006) are not included in these ratings, but are reported in Appendix A4.
2. The standard deviation across all students in each group shows how dispersed the participants' outcomes are: a smaller standard deviation on a given measure would indicate that participants had more similar outcomes. The standard deviations for Clements and Sarama (2006) were provided by the study authors upon WWC request.
3. The child-level sample sizes were provided by the study authors upon WWC request.
4. For Clements and Sarama (2006), the intervention group mean (61.69) equals the comparison group mean (53.22) plus the program coefficient from the author-conducted HLM analysis (8.47). For Clements and Sarama (2007), the intervention group mean equals the comparison group mean plus the mean difference.
5. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group. For Clements and Sarama (2006), the mean difference was the program coefficient from the author-conducted HLM analysis. For Clements and Sarama (2007), the mean differences were computed by the WWC and took into account pretest difference between the study groups. The resulting effect sizes may overestimate the intervention's effects when the intervention group had lower pretest scores than the comparison group and underestimate the intervention's effects when the intervention group had higher pretest scores than the comparison group.
6. For an explanation of the effect size calculation, see [Technical Details of WWC-Conducted Computations](#).
7. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups.
8. The improvement index represents the difference between the percentile rank of the average student in the intervention condition versus the percentile rank of the average student in the comparison condition. The improvement index can take on values between –50 and +50, with positive numbers denoting results favorable to the intervention group.

(continued)

## Appendix A3 Summary of study findings included in the rating for the math domain<sup>1</sup> *(continued)*

9. The level of statistical significance was reported by the study authors or, where necessary, calculated by the WWC to correct for clustering within classrooms or schools and for multiple comparisons. For an explanation about the clustering correction, see the WWC Tutorial on Mismatch. See [Technical Details of WWC-Conducted Computations](#) for the formulas the WWC used to calculate statistical significance. In the case of Clements and Sarama (2006), no corrections for clustering or multiple comparisons were needed.
10. The WWC-computed average effect sizes for each study and for the domain across studies are simple averages rounded to two decimal places. The average improvement indices are calculated from the average effect sizes.
11. In the case of Clements & Sarama (2007), corrections for clustering and multiple comparisons were needed, so the significance levels differ from those reported in the original study.

## Appendix A4 Summary of findings for comparisons between *Building Blocks for Math* and *Pre-K Mathematics* for the math domain<sup>1</sup>

Outcome measure	Study sample	Sample size (teachers/children) <sup>3</sup>	Authors' findings from the study		WWC calculations			
			Mean outcome (standard deviation <sup>2</sup> )		Mean difference <sup>5</sup> ( <i>Building Blocks for Math</i> – <i>Pre-K Math</i> )	Effect size <sup>6</sup>	Statistical significance <sup>7</sup> (at $\alpha = 0.05$ )	Improvement index <sup>8</sup>
			<i>Building Blocks for Math</i> group <sup>4</sup>	<i>Pre-K Math</i> group <sup>4</sup>				
<b>Clements &amp; Sarama, 2006 (randomized controlled trial)<sup>9</sup></b>								
Early Mathematics Assessment	Preschool children	21/152	57.32 (7.46)	53.77 (6.53)	3.55	0.49	Statistically significant	+19
<b>Domain average<sup>10</sup> for math</b>						0.49	Statistically significant	+19

BBA = Building Blocks Assessment of Early Mathematics, PreK–K

1. This appendix presents findings for the head-to-head comparison of *Building Blocks for Math* and *Pre-K Mathematics* for a measure that falls in the math domain. *Pre-K Mathematics* was implemented in conjunction with an additional component (DLM Express software). Comparisons of *Building Blocks for Math* and the business-as-usual comparison group were used for rating purposes and are presented in Appendix A3.
2. The standard deviation across all students in each group shows how dispersed the participants' outcomes are: a smaller standard deviation on a given measure would indicate that participants had more similar outcomes. The standard deviations were provided by the study authors upon WWC request.
3. The child-level sample size was provided by the study authors upon WWC request.
4. The *Building Blocks for Math* group mean (57.32) equals the *Pre-K Mathematics* group mean (53.77) plus the program coefficient from the author-conducted HLM analysis (3.55).
5. Positive differences and effect sizes favor the *Building Blocks for Math* group; negative differences and effect sizes favor the *Pre-K Mathematics* group. The mean difference was the program coefficient from the author-conducted HLM analysis.
6. For an explanation of the effect size calculation, see [Technical Details of WWC-Conducted Computations](#).
7. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups.
8. The improvement index represents the difference between the percentile rank of the average student in the *Building Blocks for Math* condition versus the percentile rank of the average student in the *Pre-K Mathematics* condition. The improvement index can take on values between –50 and +50, with positive numbers denoting results favorable to the *Building Blocks for Math* group.
9. The level of statistical significance was reported by the study authors or, where necessary, calculated by the WWC to correct for clustering within classrooms or schools and for multiple comparisons. For an explanation about the clustering correction, see the [WWC Tutorial on Mismatch](#). See [Technical Details of WWC-Conducted Computations](#) for the formulas the WWC used to calculate statistical significance. In the case of Clements and Sarama (2006), no corrections for clustering or multiple comparisons were needed.
10. This row provides the study average, which in this instance, is also the domain average. The WWC-computed domain average effect size is a simple average rounded to two decimal places. The domain improvement index is calculated from the average effect size.

(continued)

## Appendix A5 *Building Blocks for Math* rating for the math domain

The WWC rates an intervention's effects in a given outcome domain as positive, potentially positive, mixed, no discernible effects, potentially negative, or negative.<sup>1</sup>

For the outcome domain of math, the WWC rated *Building Blocks for Math* as having positive effects. The remaining ratings (potentially positive effects, mixed effects, no discernible effects, potentially negative effects, and negative effects) were not considered because *Building Blocks for Math* was assigned the highest applicable rating.

### Rating received

**Positive effects:** Strong evidence of a positive effect with no overriding contrary evidence.

- Criterion 1: Two or more studies showing statistically significant *positive* effects, at least one of which met WWC evidence standards for a strong design.

**Met.** Two studies of *Building Blocks for Math* showed statistically significant positive effects and both studies met WWC evidence standards for a strong design.

### AND

- Criterion 2: No studies showing statistically significant or substantively important *negative* effects.

**Met.** No studies showed statistically significant or substantively important negative effects.

1. For rating purposes, the WWC considers the statistical significance of individual outcomes and the domain-level effect. The WWC also considers the size of the domain-level effect for ratings of potentially positive or potentially negative effects. See the [WWC Intervention Rating Scheme](#) for a complete description.

## Appendix A6 Extent of evidence by domain

Outcome domain	Number of studies	Sample size		Extent of evidence <sup>1</sup>
		Schools	Classrooms/children	
Oral language	0	0	0	na
Print knowledge	0	0	0	na
Phonological processing	0	0	0	na
Early reading/writing	0	0	0	na
Cognition	0	0	0	na
Math	2	4+	32/255	Small

na = not applicable/not studied

1. A rating of “moderate to large” requires at least two studies and two schools across studies in one domain, and a total sample size across studies of at least 350 students or 14 classrooms. Otherwise, the rating is “small.”