

# Appendix

## Appendix A1.1 Study characteristics: RMC Research Corporation, 2003 (randomized controlled trial)

Characteristic	Description
<b>Study citation</b>	RMC Research Corporation. (2003) <i>Ready, Set, Leap! program: Newark prekindergarten study 2002-2003 final report</i> . Retrieved from LeapFrog SchoolHouse website: <a href="http://www.leapfrogschoolhouse.com/content/research/RMC_RSLreport.pdf">http://www.leapfrogschoolhouse.com/content/research/RMC_RSLreport.pdf</a>
<b>Participants</b>	Seventeen schools were randomly assigned either to an intervention (N = 8) or a comparison (N = 9) group. The study began with 308 inner-city, low-income preschool children enrolled in 34 classrooms in these 17 schools. The researchers excluded seven of the 34 classrooms because they included only children with moderate to severe disabilities. An additional 20 children were lost to attrition, resulting in a final sample of 254 children. <sup>1</sup> The final sample included 129 children in the intervention group and 125 children in the comparison group. At posttest, the mean age of the children in the intervention group was 4.5 years; 57% were female; and 51% were African-American, 42% Hispanic, 5% Caucasian, and 2% Asian or other race/ethnicity. At posttest, the mean age of the children in the comparison group was 4.5 years; 53% were female; and 37% were African-American, 32% Hispanic, 24% Caucasian, and 7% Asian or other race/ethnicity. The difference in the proportion of minority students was statistically significant.
<b>Setting</b>	The study took place in 17 preschools in Newark, New Jersey.
<b>Intervention</b>	<i>Ready, Set, Leap!</i> <sup>®</sup> is a prekindergarten curriculum that focuses on developing early reading skills such as phonemic awareness, letter knowledge, and letter–sound correspondence. For the study, the curriculum was integrated into the existing High/Scope framework. According to the developer’s website ( <a href="http://www.highscope.org">www.highscope.org</a> ), High/Scope is a flexible framework for setting up and managing a preschool classroom. “Active learning” is a central tenet of the approach in which children are encouraged to learn through direct, hands-on experiences. Adults support that learning through scaffolding and interaction, using techniques such as focusing on children’s strengths and problem solving. As in the comparison classrooms, letter names were taught daily. Information on implementation in the intervention classrooms was gathered through several methods, such as classroom observations, and the authors concluded that the curriculum was not fully implemented in all classrooms.
<b>Comparison</b>	The comparison classrooms also used the High/Scope framework. As in the intervention classrooms, letter names were taught daily. Although classroom observations were collected of the comparison classrooms, no information about the implementation of the High/Scope curriculum in these classrooms was provided.
<b>Primary outcomes and measurement</b>	The primary outcome domains assessed were children’s oral language, print knowledge, phonological processing, and early reading/writing. Oral language was assessed with the Peabody Picture Vocabulary Test-III (PPVT-III). Print knowledge was assessed with the Dynamic Indicators of Basic Early Literacy Skills (DIBELS) Letter Naming Fluency subtest and the Woodcock-Johnson III (WJ III) Letter-Word Identification subtest. Phonological processing was assessed with the DIBELS Initial Sound Fluency subtest, the WJ III Sound Awareness-Rhyming subtest, and the Comprehensive Test of Phonological Processing (CTOPP) Blending Words subtest. Early reading/writing was assessed with the WJ III Passage Comprehension subtest. For a more detailed description of these outcome measures, see Appendix A2.1-2.4. <sup>2</sup>
<b>Staff/teacher training</b>	The intervention group teachers received three days of training on the <i>Ready, Set, Leap!</i> <sup>®</sup> curriculum over the course of the year.

1. Information about the numbers of children included in the classrooms who were excluded from the analysis was provided by the study authors upon request from the WWC.
2. The authors also developed a phonological awareness composite (based on average raw scores from the tests of initial sound fluency, blending, and rhyming) and a letter identification composite (based on average raw scores from tests of letter-word identification, passage comprehension, and letter naming fluency). In this intervention report, the WWC reports the individual measures, rather than the composite. 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: PCER Research Corporation, 2008 (randomized controlled trial)

Characteristic	Description
<b>Study citation</b>	Preschool Curriculum Evaluation Research (PCER) Consortium (2008). <i>Effects of preschool curriculum programs on school readiness</i> (NCER 2008-2009). Washington, DC: National Center for Education Research, Institute for Education Sciences, US Department of Education. Washington, DC: US Government Printing Office.
<b>Participants</b>	Thirty-nine classrooms from 21 schools were randomly assigned either to an intervention (21 classrooms from 11 schools) or a comparison (18 classrooms from 10 schools) group. Before random assignment, schools that had similar characteristics, such as teachers' experience, school location, or a score on a state report card, were placed in blocks. Random assignment then was conducted within each block. The study began with 286 preschool children (149 in the intervention group and 137 in the comparison group). The response rate was 96% in the fall and 92% in the spring of the prekindergarten year. At pretest, the mean age of the children in the intervention group was 4.5 years; 52% were male; 82% were African-American and 18% Hispanic. At pretest, the mean age of the children in the comparison group was 4.5 years; 57% were male; 75% were African-American and 23% Hispanic. Differences between the intervention and comparison groups on these characteristics were not statistically significant.
<b>Setting</b>	The study took place in 21 full-day preschools in an urban area of New Jersey. All of the preschools in the study had National Association for the Education of Young Children (NAEYC) certification.
<b>Intervention</b>	<i>Ready, Set, Leap!</i> <sup>®</sup> is a prekindergarten curriculum that focuses on developing early reading skills, such as phonemic awareness, letter knowledge, and letter-sound correspondence. The curriculum is structured around 9 thematic units, each with 120 lessons plans for large- and small-group instruction. The research team used multiple sources to assess implementation of the curriculum, including coaching visits, site coordinator ratings, and class observations. This information was used to create a four-point scale of fidelity from "Not at all" (0) to "High" (3). The treatment classrooms received an average rating of 1.9.
<b>Comparison</b>	The comparison classrooms used what the study authors described as a "High/Scope approach." According to the developer's website ( <a href="http://www.highscope.org">www.highscope.org</a> ), High/Scope is a flexible framework for setting up and managing a preschool classroom. "Active learning" is a central tenet of the approach in which children are encouraged to learn through direct, hands-on experiences. Adults support that learning through scaffolding and interaction, using techniques such as focusing on children's strengths and problem solving. Implementation of the comparison curriculum was evaluated using the same procedures described for the intervention classrooms. The comparison classrooms received an average rating of 2.0 (out of 3).
<b>Primary outcomes and measurement</b>	The primary outcome domains assessed were the children's oral language, print knowledge, phonological processing, and math. Oral language was assessed with the Peabody Picture Vocabulary Test-III (PPVT-III) and the Test of Language Development-Primary III (TOLD-P:3) Grammatical Understanding subtest. Print knowledge was assessed with the Test of Early Reading Ability-III (TERA-3), Woodcock-Johnson III (WJ III) Letter-Word Identification subtest, and the WJ III Spelling subtest. Phonological processing was assessed with the Preschool Comprehensive Test of Phonological and Print Processing (Pre-CTOPPP) Elision subtest. Math was assessed with the WJ III Applied Problems subtest, the Child Math Assessment-Abbreviated (CMA-A), and the Building Blocks, Shape Composition task. For a more detailed description of these outcome measures, see Appendix A2.1-2.5.
<b>Staff/teacher training</b>	The intervention group teachers received four full days of training on the <i>Ready, Set, Leap!</i> <sup>®</sup> curriculum over the course of the year.

## Appendix A2.1 Outcome measures for the oral language domain

Outcome measure	Description
<b>Peabody Picture Vocabulary Test-III (PPVT-III)</b>	A standardized measure of children's receptive vocabulary that requires children to identify pictures that correspond to spoken words (cited in RMC Research Corporation, 2003).
<b>Test of Language Development-Primary III (TOLD-P:3) Grammatical Understanding subtest</b>	A standardized measure of children's ability to comprehend the meaning of sentences by selecting pictures that most accurately represents the sentence (cited in PCER Consortium, 2008).

## Appendix A2.2 Outcome measures for the print knowledge domain

Outcome measure	Description
<b>Woodcock-Johnson III (WJ III) Letter-Word Identification subtest</b>	A standardized measure of children's ability to name printed letters and words (cited in RMC Research Corporation, 2003).
<b>Dynamic Indicators of Basic Early Literacy Skills (DIBELS) Letter Naming Fluency subtest</b>	A timed standardized measure to assess children's ability to name printed upper- and lowercase letters in about one minute (cited in RMC Research Corporation, 2003).
<b>Test of Early Reading Ability III (TERA-3)</b>	A standardized measure of children's developing reading skills with three subtests: alphabet, conventions, and meaning (cited in PCER Consortium, 2008). <sup>1</sup>
<b>WJ III Spelling subtest</b>	A standardized measure that assesses children's prewriting skills, such as drawing lines, tracing, and writing letters (cited in PCER Consortium, 2008).

## Appendix A2.3 Outcome measures for the phonological processing domain

Outcome measure	Description
<b>Comprehensive Test of Phonological Processing (CTOPP) Blending Words subtest</b>	A standardized measure of children's ability to blend orally presented sounds to form words but, unlike the Pre-CTOPPP, does not use pictures when administered (cited in RMC Research Corporation, 2003, and PCER Consortium, 2008).

1. By name, this measure sounds like it should be captured under the early reading/writing domain; however, the description of the measure identifies constructs that are pertinent to print knowledge, such as knowing the alphabet, understanding print conventions, and environmental print.

(continued)

### Appendix A2.3 Outcome measures for the phonological processing domain *(continued)*

Outcome measure	Description
<b>DIBELS Initial Sound Fluency subtest</b>	A timed standardized measure to assess children's ability to recognize and produce the initial sound in an orally presented word in about one minute (cited in RMC Research Corporation, 2003).
<b>WJ III Sound Awareness-Rhyming subtest</b>	A standardized measure of children's ability to identify word sounds and rhymes when presented orally (cited in RMC Research Corporation, 2003).
<b>Preschool Comprehensive Test of Phonological and Print Processing (Pre-CTOPPP), Elision subtest</b>	A measure 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 (cited in PCER Consortium, 2008).

### Appendix A2.4 Outcome measures for the early reading/writing domain

Outcome measure	Description
<b>WJ III Passage Comprehension subtest</b>	A standardized measure of children's listening and reading comprehension skills that uses a cloze procedure (cited in RMC Research Corporation, 2003).

### Appendix A2.5 Outcome measures for the math domain

Outcome measure	Description
<b>WJ III Applied Problems subtest</b>	A standardized measure of children's ability to solve numerical and spatial problems, presented verbally with accompanying pictures of objects (cited in PCER Consortium, 2008).
<b>Child Math Assessment-Abbreviated (CMA-A) Composite Score</b>	The average of four subscales: (1) solving addition and subtraction problems using visible objects, (2) constructing a set of objects equal in number to a given set, (3) recognizing shapes, and (4) copying a pattern using objects that vary in color and identity from the model pattern (cited in PCER Consortium, 2008).
<b>Building Blocks, Shape Composition task</b>	Modified for PCER from the Building Blocks assessment tools. Children use blocks to fill in a puzzle and are assessed on whether they fill the puzzle without gaps or hangovers (cited in PCER Consortium, 2008).

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

Outcome measure	Study sample	Sample size (schools/ students)	Authors' findings from the study		WWC calculations			
			Mean outcome (standard deviation) <sup>2</sup>		Mean difference <sup>3</sup> (Ready, Set, Leap! <sup>®</sup> – comparison)	Effect size <sup>4</sup>	Statistical significance <sup>5</sup> (at $\alpha = 0.05$ )	Improvement index <sup>6</sup>
			Ready, Set, Leap! <sup>®</sup> group	Comparison group				
<b>RMC Research Corporation, 2003 (randomized controlled trial)<sup>7</sup></b>								
PPVT-III	4 year olds	17/254	56.73 (16.13)	56.59 (13.82)	0.14	0.01	ns	0
<b>Average for oral language (RMC Research Corporation, 2003)<sup>8</sup></b>						<b>0.01</b>	<b>ns</b>	<b>0</b>
<b>PCER Consortium, 2008 (randomized controlled trial)<sup>7</sup></b>								
PPVT-III	Preschoolers	21/260	nr	nr	nr	0.15	ns	+6
TOLD-P:3 Grammatical Understanding subtest	Preschoolers	21/258	nr	nr	nr	-0.11	ns	-4
<b>Average for oral language (PCER Consortium, 2008)<sup>8</sup></b>						<b>0.02</b>	<b>ns</b>	<b>+1</b>
<b>Domain average for oral language across all studies<sup>8</sup></b>						<b>0.02</b>	<b>na</b>	<b>+1</b>

na = not applicable

nr = not reported

ns = not statistically significant

PPVT-III = Peabody Picture Vocabulary Test-III

TOLD-P:3 = Test of Language Development Primary, Third Edition

1. This appendix reports findings considered for the effectiveness rating and the average improvement indices for the oral language domain. Follow-up findings from the PCER Consortium (2008) study are not included in these ratings, but are reported in Appendix A4.1.
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 the RMC Research Corporation (2003) study were provided by the study authors upon request from the WWC.
3. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group. The means reported for RMC Research Corporation (2003) are regression adjusted.
4. For an explanation of the effect size calculation, see [Technical Details of WWC-Conducted Computations](#). The effect sizes for the RMC Research Corporation (2003) study are slightly different than those reported by the study authors due to a difference in the formula used by WWC to compute effect sizes. In the case of PCER Consortium (2008), the WWC used the effect sizes reported by the study authors.
5. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups.
6. The improvement index represents the difference between the percentile rank of the average student in the intervention condition and that 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.
7. 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](#). For the formulas the WWC used to calculate statistical significance, see [Technical Details of WWC-Conducted Computations](#). In the case of RMC Research Corporation (2003), the study authors corrected for clustering within classrooms, and the WWC did not apply additional corrections. In the case of PCER Consortium (2008), a clustering correction was not needed.
8. 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.

## Appendix A3.2 Summary of study findings included in the rating for the print knowledge domain<sup>1</sup>

Outcome measure	Study sample	Sample size (schools/ students)	Authors' findings from the study		WWC calculations			
			Mean outcome (standard deviation) <sup>2</sup>		Mean difference <sup>3</sup> ( <i>Ready, Set, Leap!</i> <sup>®</sup> – comparison)	Effect size <sup>4</sup>	Statistical significance <sup>5</sup> (at $\alpha = 0.05$ )	Improvement index <sup>6</sup>
			<i>Ready, Set, Leap!</i> <sup>®</sup> group	Comparison group				
<b>RMC Research Corporation, 2003 (randomized controlled trial)<sup>7</sup></b>								
WJ III Letter Word Identification subtest	4 year olds	17/254	13.59 (5.70)	12.94 (5.06)	0.65	0.12	ns	+5
DIBELS Letter Naming Fluency subtest	4 year olds	17/254	23.63 (14.94)	24.76 (14.72)	-1.13	-0.08	ns	-3
<b>Average for print knowledge (RMC Research Corporation, 2003)<sup>8</sup></b>						<b>0.02</b>	<b>ns</b>	<b>+1</b>
<b>PCER Consortium, 2008 (randomized controlled trial)<sup>7</sup></b>								
TERA	Preschoolers	21/257	nr	nr	nr	0.08	ns	+3
WJ III Letter-word identification subtest	Preschoolers	21/257	nr	nr	nr	0.01	ns	0
WJ III Spelling subtest	Preschoolers	21/236	nr	nr	nr	0.02	ns	+8
<b>Average for print knowledge (PCER Consortium, 2008)<sup>8</sup></b>						<b>0.10</b>	<b>ns</b>	<b>+4</b>
<b>Domain average for print knowledge across all studies<sup>8</sup></b>						<b>0.06</b>	<b>na</b>	<b>+2</b>

na = not applicable    nr = not reported    ns = not statistically significant

WJ III = Woodcock-Johnson III

DIBELS = Dynamic Indicators of Basic Early Literacy Skills

TERA= Test of Early Reading Ability

1. This appendix reports findings considered for the effectiveness rating and the average improvement indices for the print knowledge domain. Follow-up findings from the PCER Consortium (2008) study are not included in these ratings, but are reported in Appendix A4.2.
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 the RMC Research Corporation (2003) study were provided by the study authors upon the WWC request.
3. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group. The means reported for RMC Research Corporation (2003) are regression adjusted.
4. For an explanation of the effect size calculation, see [Technical Details of WWC-Conducted Computations](#). The effect sizes for the RMC Research Corporation (2003) study are slightly different than those reported by the study authors due to a difference in the formula used by WWC to compute effect sizes. In the case of PCER Consortium (2008), the WWC used the effect sizes reported by the study authors.
5. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups.
6. The improvement index represents the difference between the percentile rank of the average student in the intervention condition and that 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.
7. 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](#). For the formulas the WWC used to calculate statistical significance, see [Technical Details of WWC-Conducted Computations](#). In the case of RMC Research Corporation (2003), the study authors corrected for clustering within classrooms, and the WWC did not apply additional corrections. In the case of PCER Consortium (2008), a clustering correction was not needed.
8. 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.

## Appendix A3.3 Summary of study findings included in the rating for the phonological processing domain<sup>1</sup>

Outcome measure	Study sample	Sample size (schools/ students)	Authors' findings from the study		WWC calculations			
			Mean outcome (standard deviation) <sup>2</sup>		Mean difference <sup>3</sup> (Ready, Set, Leap! <sup>®</sup> – comparison)	Effect size <sup>4</sup>	Statistical significance <sup>5</sup> (at $\alpha = 0.05$ )	Improvement index <sup>6</sup>
			Ready, Set, Leap! <sup>®</sup> group	Comparison group				
<b>RMC Research Corporation, 2003 (randomized controlled trial)<sup>7</sup></b>								
CTOPP Blending Words subtest	4 year olds	17/254	4.24 (4.17)	3.18 (3.38)	1.06	0.28	ns	+11
DIBELS Initial Sound Fluency subtest	4 year olds	17/254	11.03 (8.24)	9.58 (6.48)	1.45	0.20	ns	+8
WJ III Sound Awareness-Rhyming subtest	4 year olds	17/254	5.49 (4.10)	4.92 (3.72)	0.57	0.15	ns	+6
<b>Average for phonological processing (RMC Research Corporation, 2003)<sup>8</sup></b>						<b>0.21</b>	<b>ns</b>	<b>+8</b>
<b>PCER Consortium, 2008 (randomized controlled trial)<sup>7</sup></b>								
Pre-CTOPPP Elision subtest	Preschoolers	21/262	nr	nr	nr	-0.09	ns	-4
<b>Average for phonological processing (PCER Consortium, 2008)<sup>8</sup></b>						<b>-0.09</b>	<b>ns</b>	<b>-4</b>
<b>Domain average for phonological processing across all studies<sup>8</sup></b>						<b>0.06</b>	<b>na</b>	<b>+2</b>

na = not applicable    nr = not reported    ns = not statistically significant

CTOPP = Comprehensive Test of Phonological Processing

DIBELS = Dynamic Indicators of Basic Early Literacy Skills

WJ III = Woodcock-Johnson III

Pre-CTOPPP = Preschool Comprehensive Test of Phonological and Print Processing

1. This appendix reports findings considered for the effectiveness rating and the average improvement indices for the phonological processing domain.
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 the RMC Research Corporation (2003) study were provided by the study authors upon the WWC request.
3. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group. The means reported for RMC Research Corporation (2003) are regression adjusted.
4. For an explanation of the effect size calculation, see [Technical Details of WWC-Conducted Computations](#). The effect sizes for the RMC Research Corporation (2003) study are slightly different than those reported by the study authors due to a difference in the formula used by WWC to compute effect sizes. In the case of PCER Consortium (2008), the WWC used the effect sizes reported by the study authors.
5. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups.
6. The improvement index represents the difference between the percentile rank of the average student in the intervention condition and that 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.
7. 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](#). For the formulas the WWC used to calculate statistical significance, see [Technical Details of WWC-Conducted Computations](#). In the case of RMC Research Corporation (2003), the study authors corrected for clustering within classrooms, and the WWC did not apply additional corrections. In the case of PCER Consortium (2008), a clustering correction was not needed.
8. 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.

## Appendix A3.4 Summary of study findings included in the rating for the early reading/writing domain<sup>1</sup>

Outcome measure	Study sample	Sample size (schools/students)	Authors' findings from the study		WWC calculations			
			Mean outcome (standard deviation) <sup>2</sup>		Mean difference <sup>3</sup> (Ready, Set, Leap! <sup>®</sup> – comparison)	Effect size <sup>4</sup>	Statistical significance <sup>5</sup> (at $\alpha = 0.05$ )	Improvement index <sup>6</sup>
			Ready, Set, Leap! <sup>®</sup> group	Comparison group				
<b>RMC Research Corporation, 2003 (randomized controlled trial)<sup>7</sup></b>								
WJ III Passage Comprehension subtest	4 year olds	17/254	5.89 (2.43)	5.69 (2.24)	0.20	0.09	ns	+3
<b>Domain average for early reading/writing<sup>8</sup></b>						<b>0.09</b>	<b>na</b>	<b>+3</b>

na = not applicable

ns = not statistically significant

WJ III = Woodcock-Johnson III

1. This appendix reports findings considered for the effectiveness rating and the average improvement indices for the early reading/writing domain.
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 the RMC Research Corporation (2003) study were provided by the study authors upon the WWC request.
3. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group. The means reported for RMC Research Corporation (2003) are regression adjusted.
4. For an explanation of the effect size calculation, see [Technical Details of WWC-Conducted Computations](#). The effect sizes for the RMC Research Corporation (2003) study are slightly different than those reported by the study authors due to a difference in the formula used by WWC to compute effect sizes. In the case of PCER Consortium (2008), the WWC used the effect sizes reported by the study authors.
5. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups.
6. The improvement index represents the difference between the percentile rank of the average student in the intervention condition and that 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.
7. 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](#). For the formulas the WWC used to calculate statistical significance, see [Technical Details of WWC-Conducted Computations](#). In the case of RMC Research Corporation (2003), the study authors corrected for clustering within classrooms, and the WWC did not apply additional corrections.
8. This row provides the study average, which in this instance is also the domain average. The WWC-computed domain effect size is a simple average rounded to two decimal places. The average improvement indices are calculated from the average effect sizes.



## Appendix A3.5 Summary of study findings included in the rating for the math domain<sup>1</sup>

Outcome measure	Study sample	Sample size (schools/students)	Authors' findings from the study		WWC calculations			
			Mean outcome (standard deviation) <sup>2</sup>		Mean difference <sup>3</sup> (Ready, Set, Leap! <sup>®</sup> – comparison)	Effect size <sup>4</sup>	Statistical significance <sup>5</sup> (at $\alpha = 0.05$ )	Improvement index <sup>6</sup>
			Ready, Set, Leap! <sup>®</sup> group	Comparison group				
<b>PCER Consortium, 2008 (randomized controlled trial)<sup>7</sup></b>								
WJ III Applied Problems	Preschoolers	21/251	nr	nr	nr	0.04	ns	+2
CMA-A Composite	Preschoolers	21/274	nr	nr	nr	-0.24	ns	-9
Shape Composition	Preschoolers	21/273	nr	nr	nr	0.08	ns	+3
<b>Domain average for math<sup>8</sup></b>						<b>-0.04</b>	<b>na</b>	<b>-2</b>

na = not applicable

nr = not reported

ns = not statistically significant

WJ III = Woodcock-Johnson III

CMA-A = Child Math Assessment-Abbreviated

1. This appendix reports findings considered for the effectiveness rating and the average improvement indices for the math domain. Follow-up findings from the PCER Consortium (2008) study are not included in these ratings, but are reported in Appendix A4.3.
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.
3. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group.
4. For an explanation of the effect size calculation, see [Technical Details of WWC-Conducted Computations](#). In the case of PCER Consortium (2008), the WWC used the effect sizes reported by the study authors.
5. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups.
6. The improvement index represents the difference between the percentile rank of the average student in the intervention condition and that 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.
7. 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](#). For the formulas the WWC used to calculate statistical significance, see [Technical Details of WWC-Conducted Computations](#). In the case of PCER Consortium (2008), a correction for clustering was not needed, but a correction for multiple comparisons was used, so the significance levels may differ from those reported in the original study.
8. This row provides the study average, which in this instance is also the domain average. The WWC-computed domain effect size is a simple average rounded to two decimal places. The average improvement indices are calculated from the average effect sizes.

## Appendix A4.1 Summary of kindergarten follow-up findings for the oral language domain<sup>1</sup>

Outcome measure	Study sample	Sample size (schools/students)	Authors' findings from the study		WWC calculations			
			Mean outcome (standard deviation) <sup>2</sup>		Mean difference <sup>3</sup> (Ready, Set, Leap! <sup>®</sup> – comparison)	Effect size <sup>4</sup>	Statistical significance <sup>5</sup> (at $\alpha = 0.05$ )	Improvement index <sup>6</sup>
			Ready, Set, Leap! <sup>®</sup> group	Comparison group				
<b>PCER Consortium, 2008 (randomized controlled trial)<sup>7</sup></b>								
PPVT-III	Kindergarteners	94/240	nr	nr	nr	–0.02	ns	–1
TOLD-P:3 Grammatical Understanding subtest	Kindergarteners	94/247	nr	nr	nr	–0.03	ns	–1

ns = not statistically significant

nr = not reported

PPVT-III = Peabody Picture Vocabulary Test-III

TOLD-P:3 = Test of Language Development Primary, Third Edition

1. This appendix presents follow-up findings from spring of the kindergarten year for measures that fall in the oral language domain. Posttest preschool scores were used for rating purposes and are presented in Appendix A3.1.
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.
3. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group.
4. For an explanation of the effect size calculation, see [Technical Details of WWC-Conducted Computations](#). In the case of PCER Consortium (2008), the WWC used the effect sizes reported by the study authors.
5. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups.
6. The improvement index represents the difference between the percentile rank of the average student in the intervention condition and that 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.
7. 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 (corrections for multiple comparisons were not done for findings not included in the overall intervention rating). For an explanation about the clustering correction, see the [WWC Tutorial on Mismatch](#). For the formulas the WWC used to calculate statistical significance, see [Technical Details of WWC-Conducted Computations](#). In the case of PCER Consortium (2008), no correction for clustering was needed.

## Appendix A4.2 Summary of kindergarten follow-up findings for the print knowledge domain<sup>1</sup>

Outcome measure	Study sample	Sample size (schools/students)	Authors' findings from the study					
			Mean outcome (standard deviation) <sup>2</sup>		WWC calculations			
			<i>Ready, Set, Leap!</i> <sup>®</sup> group	Comparison group	Mean difference <sup>3</sup> ( <i>Ready, Set, Leap!</i> <sup>®</sup> – comparison)	Effect size <sup>4</sup>	Statistical significance <sup>5</sup> (at $\alpha = 0.05$ )	Improvement index <sup>6</sup>
<b>PCER Consortium, 2008 (randomized controlled trial)<sup>7</sup></b>								
TERA	Kindergarteners	94/234	nr	nr	nr	0.01	ns	0
WJ III Letter-word identification subtest	Kindergarteners	94/243	nr	nr	nr	–0.12	ns	–5
WJ III Spelling subtest	Kindergarteners	94/223	nr	nr	nr	0.04	ns	+2

nr = not reported

ns = not statistically significant

TERA = Test of Early Reading Ability

WJ III = Woodcock-Johnson III

1. This appendix presents follow-up findings from spring of the kindergarten year for measures that fall in the print knowledge domain. Posttest preschool scores were used for rating purposes and are presented in Appendix A3.2.
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.
3. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group.
4. For an explanation of the effect size calculation, see [Technical Details of WWC-Conducted Computations](#). In the case of PCER Consortium (2008), the WWC used the effect sizes reported by the study authors.
5. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups.
6. The improvement index represents the difference between the percentile rank of the average student in the intervention condition and that 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.
7. 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 (corrections for multiple comparisons were not done for findings not included in the overall intervention rating). For an explanation about the clustering correction, see the [WWC Tutorial on Mismatch](#). For the formulas the WWC used to calculate statistical significance, see [Technical Details of WWC-Conducted Computations](#). In the case of PCER Consortium (2008), no correction for clustering was needed.

## Appendix A4.3 Summary of kindergarten follow-up findings for the math domain<sup>1</sup>

Outcome measure	Study sample	Sample size (schools/ students)	Authors' findings from the study		WWC calculations			
			Mean outcome (standard deviation) <sup>2</sup>		Mean difference <sup>3</sup> (Ready, Set, Leap! <sup>®</sup> – comparison)	Effect size <sup>4</sup>	Statistical significance <sup>5</sup> (at $\alpha = 0.05$ )	Improvement index <sup>6</sup>
			Ready, Set, Leap! <sup>®</sup> group	Comparison group				
<b>PCER Consortium, 2008 (randomized controlled trial)<sup>7</sup></b>								
WJ III Applied Problems	Kindergarteners	94/243	nr	nr	nr	0.00	ns	0
CMA-A Composite	Kindergarteners	94/249	nr	nr	nr	–0.10	ns	–4
Shape Composition	Kindergarteners	94/247	nr	nr	nr	0.03	ns	+1

nr = not reported

ns = not statistically significant

WJ III = Woodcock-Johnson III

CMA-A = Child Math Assessment Abbreviated

1. This appendix presents follow-up findings from spring of the kindergarten year for measures that fall in the math domain. Posttest preschool scores were used for rating purposes and are presented in Appendix A3.5.
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.
3. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group.
4. For an explanation of the effect size calculation, see [Technical Details of WWC-Conducted Computations](#). In the case of PCER Consortium (2008), the WWC used the effect sizes reported by the study authors.
5. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups.
6. The improvement index represents the difference between the percentile rank of the average student in the intervention condition and that 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.
7. 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 (corrections for multiple comparisons were not done for findings not included in the overall intervention rating). For an explanation about the clustering correction, see the [WWC Tutorial on Mismatch](#). For the formulas the WWC used to calculate statistical significance, see [Technical Details of WWC-Conducted Computations](#). In the case of PCER Consortium (2008), no correction for clustering was needed.

## Appendix A5.1 *Ready, Set, Leap!*<sup>®</sup> rating for the oral language 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 oral language, the WWC rated *Ready, Set, Leap!*<sup>®</sup> as having no discernible effects. The remaining ratings (potentially negative or negative) were not considered, as *Ready, Set, Leap!*<sup>®</sup> was assigned the highest applicable rating.

### Rating received

**No discernible effects:** No affirmative evidence of effects.

- Criterion 1: None of the studies shows a statistically significant or substantively important effect, either *positive* or *negative*.

**Met.** *Ready, Set, Leap!*<sup>®</sup> had no studies that showed statistically significant or substantively important effects on oral language.

### Other ratings considered

**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.

**Not met.** *Ready, Set, Leap!*<sup>®</sup> had no studies that showed statistically significant or substantively important positive effects.

#### AND

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

**Met.** *Ready, Set, Leap!*<sup>®</sup> had no studies that showed statistically significant or substantively important negative effects.

**Potentially positive effects:** Evidence of a positive effect with no overriding contrary evidence.

- Criterion 1: At least one study showing a statistically significant or substantively important *positive* effect.

**Not met.** *Ready, Set, Leap!*<sup>®</sup> had no studies that showed statistically significant or substantively important positive effects.

#### AND

- Criterion 2: No studies showing a statistically significant or substantively important *negative* effect and fewer or the same number of studies showing *indeterminate* effects than showing statistically significant or substantively important *positive* effects.

**Met.** *Ready, Set, Leap!*<sup>®</sup> had no studies that showed statistically significant or substantively important negative effects.

**Mixed effects:** Evidence of inconsistent effects as demonstrated through either of the following criteria.

- Criterion 1: At least one study showing a statistically significant or substantively important *positive* effect, and at least one study showing 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.

**Not met.** *Ready, Set, Leap!*<sup>®</sup> had no studies that showed statistically significant or substantively important effects, either positive or negative.

#### OR

- Criterion 2: At least one study showing a statistically significant or substantively important effect, and more studies showing an *indeterminate* effect than showing a statistically significant or substantively important effect.

**Not met.** *Ready, Set, Leap!*<sup>®</sup> had no studies that showed statistically significant or substantively important effects, either positive or negative.

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. For a complete description, see the [WWC Intervention Rating Scheme](#).

## Appendix A5.2 *Ready, Set, Leap!*<sup>®</sup> rating for the print knowledge 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 print knowledge, the WWC rated *Ready, Set, Leap!*<sup>®</sup> as having no discernible effects. The remaining ratings (potentially negative or negative) were not considered, as *Ready, Set, Leap!*<sup>®</sup> was assigned the highest applicable rating.

### Rating received

**No discernible effects:** No affirmative evidence of effects.

- Criterion 1: None of the studies shows a statistically significant or substantively important effect, either *positive* or *negative*.

**Met.** *Ready, Set, Leap!*<sup>®</sup> had no studies that showed statistically significant or substantively important effects on print knowledge.

### Other ratings considered

**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.

**Not met.** *Ready, Set, Leap!*<sup>®</sup> had no studies that showed statistically significant or substantively important positive effects.

### AND

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

**Met.** *Ready, Set, Leap!*<sup>®</sup> had no studies that showed statistically significant or substantively important negative effects.

**Potentially positive effects:** Evidence of a positive effect with no overriding contrary evidence.

- Criterion 1: At least one study showing a statistically significant or substantively important *positive* effect.

**Not met.** *Ready, Set, Leap!*<sup>®</sup> had no studies that showed statistically significant or substantively important positive effects.

### AND

- Criterion 2: No studies showing a statistically significant or substantively important *negative* effect and fewer or the same number of studies showing *indeterminate* effects than showing statistically significant or substantively important *positive* effects.

**Met.** *Ready, Set, Leap!*<sup>®</sup> had no studies that showed statistically significant or substantively important negative effects.

**Mixed effects:** Evidence of inconsistent effects as demonstrated through either of the following criteria.

- Criterion 1: At least one study showing a statistically significant or substantively important *positive* effect, and at least one study showing 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.

**Not met.** *Ready, Set, Leap!*<sup>®</sup> had no studies that showed statistically significant or substantively important effects, either positive or negative.

### OR

- Criterion 2: At least one study showing a statistically significant or substantively important effect, and more studies showing an *indeterminate* effect than showing a statistically significant or substantively important effect.

**Not met.** *Ready, Set, Leap!*<sup>®</sup> had no studies that showed statistically significant or substantively important effects, either positive or negative.

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. For a complete description, see the [WWC Intervention Rating Scheme](#).

### Appendix A5.3 *Ready, Set, Leap!*<sup>®</sup> rating for the phonological processing 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 phonological processing, the WWC rated *Ready, Set, Leap!*<sup>®</sup> as having no discernible effects. The remaining ratings (potentially negative or negative) were not considered, as *Ready, Set, Leap!*<sup>®</sup> was assigned the highest applicable rating.

#### Rating received

**No discernible effects:** No affirmative evidence of effects.

- Criterion 1: None of the studies shows a statistically significant or substantively important effect, either *positive* or *negative*.

**Met.** *Ready, Set, Leap!*<sup>®</sup> had no studies that showed statistically significant or substantively important effects on phonological processing.

#### Other ratings considered

**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.

**Not met.** *Ready, Set, Leap!*<sup>®</sup> had no studies that showed statistically significant or substantively important positive effects.

#### AND

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

**Met.** *Ready, Set, Leap!*<sup>®</sup> had no studies that showed statistically significant or substantively important negative effects.

**Potentially positive effects:** Evidence of a positive effect with no overriding contrary evidence.

- Criterion 1: At least one study showing a statistically significant or substantively important *positive* effect.

**Not met.** *Ready, Set, Leap!*<sup>®</sup> had no studies that showed statistically significant or substantively important positive effects.

#### AND

- Criterion 2: No studies showing a statistically significant or substantively important *negative* effect and fewer or the same number of studies showing *indeterminate* effects than showing statistically significant or substantively important *positive* effects.

**Met.** *Ready, Set, Leap!*<sup>®</sup> had no studies that showed statistically significant or substantively important negative effects.

**Mixed effects:** Evidence of inconsistent effects as demonstrated through either of the following criteria.

- Criterion 1: At least one study showing a statistically significant or substantively important *positive* effect, and at least one study showing 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.

**Not met.** *Ready, Set, Leap!*<sup>®</sup> had no studies that showed statistically significant or substantively important effects, either positive or negative.

#### OR

- Criterion 2: At least one study showing a statistically significant or substantively important effect, and more studies showing an *indeterminate* effect than showing a statistically significant or substantively important effect.

**Not met.** *Ready, Set, Leap!*<sup>®</sup> had no studies that showed statistically significant or substantively important effects, either positive or negative.

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. For a complete description, see the [WWC Intervention Rating Scheme](#).

## Appendix A5.4 *Ready, Set, Leap!*<sup>®</sup> rating for the early reading/writing 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 early reading/writing, the WWC rated *Ready, Set, Leap!*<sup>®</sup> as having no discernible effects. The remaining ratings (potentially negative or negative) were not considered, as *Ready, Set, Leap!*<sup>®</sup> was assigned the highest applicable rating.

### Rating received

**No discernible effects:** No affirmative evidence of effects.

- Criterion 1: None of the studies shows a statistically significant or substantively important effect, either *positive* or *negative*.

**Met.** *Ready, Set, Leap!*<sup>®</sup> had no studies that showed statistically significant or substantively important effects on early reading/writing.

### Other ratings considered

**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.

**Not met.** *Ready, Set, Leap!*<sup>®</sup> had no studies that showed statistically significant or substantively important positive effects.

### AND

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

**Met.** *Ready, Set, Leap!*<sup>®</sup> had no studies that showed statistically significant or substantively important negative effects.

**Potentially positive effects:** Evidence of a positive effect with no overriding contrary evidence.

- Criterion 1: At least one study showing a statistically significant or substantively important *positive* effect.

**Not met.** *Ready, Set, Leap!*<sup>®</sup> had no studies that showed statistically significant or substantively important positive effects.

### AND

- Criterion 2: No studies showing a statistically significant or substantively important *negative* effect and fewer or the same number of studies showing *indeterminate* effects than showing statistically significant or substantively important *positive* effects.

**Met.** *Ready, Set, Leap!*<sup>®</sup> had no studies that showed statistically significant or substantively important negative effects.

**Mixed effects:** Evidence of inconsistent effects as demonstrated through either of the following criteria.

- Criterion 1: At least one study showing a statistically significant or substantively important *positive* effect, and at least one study showing 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.

**Not met.** *Ready, Set, Leap!*<sup>®</sup> had no studies that showed statistically significant or substantively important effects, either positive or negative.

### OR

- Criterion 2: At least one study showing a statistically significant or substantively important effect, and more studies showing an *indeterminate* effect than showing a statistically significant or substantively important effect.

**Not met.** *Ready, Set, Leap!*<sup>®</sup> had no studies that showed statistically significant or substantively important effects, either positive or negative.

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. For a complete description, see the [WWC Intervention Rating Scheme](#).



## Appendix A5.5 *Ready, Set, Leap!*<sup>®</sup> 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 *Ready, Set, Leap!*<sup>®</sup> as having no discernible effects. The remaining ratings (potentially negative or negative) were not considered, as *Ready, Set, Leap!*<sup>®</sup> was assigned the highest applicable rating.

### Rating received

**No discernible effects:** No affirmative evidence of effects.

- Criterion 1: None of the studies shows a statistically significant or substantively important effect, either *positive* or *negative*.

**Met.** *Ready, Set, Leap!*<sup>®</sup> had no studies that showed statistically significant or substantively important effects on math.

### Other ratings considered

**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.

**Not met.** *Ready, Set, Leap!*<sup>®</sup> had no studies that showed statistically significant or substantively important positive effects.

#### AND

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

**Met.** *Ready, Set, Leap!*<sup>®</sup> had no studies that showed statistically significant or substantively important negative effects.

**Potentially positive effects:** Evidence of a positive effect with no overriding contrary evidence.

- Criterion 1: At least one study showing a statistically significant or substantively important *positive* effect.

**Not met.** *Ready, Set, Leap!*<sup>®</sup> had no studies that showed statistically significant or substantively important positive effects.

#### AND

- Criterion 2: No studies showing a statistically significant or substantively important *negative* effect and fewer or the same number of studies showing *indeterminate* effects than showing statistically significant or substantively important *positive* effects.

**Met.** *Ready, Set, Leap!*<sup>®</sup> had no studies that showed statistically significant or substantively important negative effects.

**Mixed effects:** Evidence of inconsistent effects as demonstrated through either of the following criteria.

- Criterion 1: At least one study showing a statistically significant or substantively important *positive* effect, and at least one study showing 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.

**Not met.** *Ready, Set, Leap!*<sup>®</sup> had no studies that showed statistically significant or substantively important effects, either positive or negative.

#### OR

- Criterion 2: At least one study showing a statistically significant or substantively important effect, and more studies showing an *indeterminate* effect than showing a statistically significant or substantively important effect.

**Not met.** *Ready, Set, Leap!*<sup>®</sup> had no studies that showed statistically significant or substantively important effects, either positive or negative.

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. For a complete description, see the [WWC Intervention Rating Scheme](#).

## Appendix A6 Extent of evidence by domain

Outcome domain	Number of studies	Sample size		Extent of evidence <sup>1</sup>
		Schools	Students	
Oral language	2	38	540	Medium to large
Print knowledge	2	38	540	Medium to large
Phonological processing	2	38	540	Medium to large
Early reading/writing	1	17	254	Small
Cognition	0	na	na	na
Math	1	21	286	Small

na = not applicable/not studied

1. A rating of “medium 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.”