What Works Clearinghouse

INSTITUTE OF EDUCATION SCIENCES

Early Childhood Education January 2009[1]

Curiosity Corner

Program description[2]

Curiosity Corner is an early childhood curriculum emphasizing children's language and literacy skills. It comprises two sets of 38 weekly thematic units—one for three-year-olds and one for four-year-olds. Program staff conduct daily lessons using sequential daily activities.

Research

One study of *Curiosity Corner* meets the What Works Clearinghouse (WWC) evidence standards, and one study meets WWC evidence standards with reservations.[3] The two studies included more than 500 preschool children from 34 preschools in Florida, Kansas, and New Jersey.[4]

Based on these two studies, the WWC considers the extent of evidence for *Curiosity Corner* to be medium to large for oral language and small for print knowledge, phonological processing, cognition, and math. No studies that meet WWC standards with or without reservations examined the effectiveness of *Curiosity Corner* in the early reading and writing domain.

Effectiveness

Curiosity Corner was found to have no discernible effects on oral language, print knowledge, phonological processing, cognition, and math.

	Oral language	*Print knowledge*	*Phonological processing*	*Early reading and writing*	*Cognition*	*Math*
Rating of effectiveness	No discernible effects	No discernible effects	No discernible effects	na	No discernible effects	No discernible effects
Improvement index[5]	Average: +2 percentile points	Average: +3 percentile points	Average: +7 percentile points	na	Average: –3 percentile points	Average: +4 percentile points
	Range: –3 to +14 percentile points	Range: +2 to +4 percentile points		na	Range: –4 to –1 percentile points	Range: 0 to +6 percentile points

na = not applicable

1. This report has been updated to include a review of two studies that have been released since 2006. A complete list and disposition of all studies reviewed is provided in the references.
2. The descriptive information for this program was obtained from a publicly-available source: the program's website (www.successforall.org/early/early_curiosity.htm, downloaded November 2008). The WWC requests developers to review the program description sections for accuracy from their perspective. Further verification of the accuracy of the descriptive information for this program is beyond the scope of this review.
3. To be eligible for the WWC's review, the Early Childhood Education intervention had to be implemented in English, in center-based settings, with children aged three to five years, or in preschool.
4. The evidence presented in this report is based on available research. Findings and conclusions may change as new research becomes available.
5. These numbers show the average and range of student-level improvement indices for all findings across the studies.

Absence of conflict of interest

The PCER Consortium (2008) study summarized in this intervention report had numerous contributors, including staff of Mathematica Policy Research, Inc. (MPR). Because the principal investigator for the WWC Early Childhood Education review is also an MPR staff member, the study was rated by Chesapeake Research Associates, 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.

Additional program information

Developer and contact

Curiosity Corner was developed and is distributed by The Success for All Foundation. Address: Success For All Foundation, Inc., 200 W. Towsontown Boulevard, Baltimore, MD 21204-5200. Email: sfainfo@successforall.org. Web: www.successforall.org/early/early_curiosity.htm. Telephone: (800) 548-4998, ext. 2372.

Scope of use

Curiosity Corner is currently implemented in more than 300 sites in 29 states. *Curiosity Corner* is used in Head Start centers, preschool classes in elementary schools, child care centers, and early childhood education centers, mostly in high-poverty neighborhoods.

Teaching

Curiosity Corner can be implemented in various early childhood settings; however, settings are typically composed of about 15 children with a teacher and an assistant. Additional teaching staff are required for children who need special care or with special needs. The program includes a teacher's manual and weekly theme guides to provide teachers with detailed instructions for lessons and supplies (themed children's books, manipulatives, and games) for the instructional activities. Teachers are also given initial training and follow-up support (workshops and in-class visits by the Success for All Foundation staff).

Curiosity Corner is organized by weekly themes. Daily activities—which proceed in a sequential order to provide children with active learning experiences—include Greetings and Readings, Clues and Questions, Rhyme Time, Learning Labs, Story Tree, Outside/Gross Motor Play, Snack Time, and Question/Reflection. Although designed to enhance the development of the whole child, the program emphasizes children's language and early literacy skills. Parents are encouraged to participate through various activities both in and out of the classroom, such as home visits, the Home Link Page, a lending library, videos, and classroom activities.

The Success for All Foundation staff provides professional development. The first year of professional development includes an initial two-day training session, additional training sessions and ongoing implementation visits, and a fall conference for *Curiosity Corner* coaches and facilitators. The second year includes a one-day refresher session, subsequent training sessions and ongoing implementation visits, and training available at experienced sites and conferences for *Curiosity Corner* coaches.

Cost[6]

Teaching materials for *Curiosity Corner* cannot be purchased without participating in training and other professional development activities. The teaching materials, costing $2,825 per class, come with a teacher's manual, 38 weekly theme guides, and more than 150 children's trade books, manipulative materials, games, and puppets. The first year of professional development costs on average $1,516 per classroom and includes initial training and follow-up support. Costs for the second year of professional development are based on the amount of time for which administrators contract.

6. The WWC requested the developer to review this information in August 2008.

Research

Three studies reviewed by the WWC investigated the effects of *Curiosity Corner*. One study is a randomized controlled trial that meets WWC evidence standards. A second study is a quasi-experimental study that meets WWC evidence standards with reservations. The remaining study does not meet WWC eligibility screens.

Meets evidence standards

One study reviewed by the WWC, PCER Consortium (2008), assesses *Curiosity Corner's* effectiveness as part of the Preschool Curriculum Evaluation Research (PCER) effort.[7] PCER Consortium (2008) used a randomized controlled trial design in which 18 preschools in Florida, Kansas, and New Jersey were randomly assigned to implement *Curiosity Corner* or to a control group. The study sample included children in 31 classrooms. Following parent consent, data were collected on 211 children. Half of the children were male, half were African-American, and 14% were reported to have a disability. Pretests were collected in the fall and posttests in the spring of the preschool year. The study investigated effects on oral language, print knowledge, phonological processing, and math. The comparison condition varied across sites and included both teacher-developed and branded curricula.

Meets evidence standards with reservations

Chambers, Chamberlain, Hurley, and Slavin (2001), the second study, investigates the effects of *Curiosity Corner* using a quasi-experimental design that meets WWC evidence standards with reservations. The study included 316 children in 16 private and public preschools (three-year-old children at private child care centers and four-year-old children at public schools from four urban, high-poverty school districts in New Jersey). More than two-thirds of the children were African-American. Pretests were collected in the fall and posttests in the spring. The authors compared oral language and cognitive outcomes for children in a *Curiosity Corner* intervention group with those for children in a comparison group that used the classroom's standard early childhood curriculum.

Extent of evidence

The WWC categorizes the extent of evidence in each domain as small or medium to large (see the What Works Clearinghouse Extent of Evidence Categorization Scheme). The extent of evidence takes into account the number of studies and the total sample size across the studies that met WWC evidence standards with or without reservations.[8] The WWC considers the extent of evidence for *Curiosity Corner* to be medium to large for oral language and small for print knowledge, phonological processing, cognition, and math. No studies that met WWC evidence standards with or without reservations addressed early reading and writing.

Effectiveness

Findings

The WWC review of interventions for early childhood education addresses children's outcomes in six domains: oral language, print knowledge, phonological processing, early reading and writing, cognition, and math. The studies included in this report cover five domains: oral language, print knowledge, phonological processing, cognition, and math. The findings below present the authors' and the WWC-calculated estimates of the size

7. Preschool Curriculum Evaluation Research (PCER 2008) evaluated a total of 14 curricula, including *Curiosity Corner*, in comparison to the respective local control conditions.
8. The Extent of Evidence Categorization was developed to tell readers how much evidence was used to determine the intervention rating, focusing on the number and size of studies. Additional factors associated with a related concept–external validity, such as the students' demographics and the types of settings in which studies took place–are not taken into account for the categorization. Information about how the extent of evidence rating was determined for *Curiosity Corner* is in Appendix A6.

Effectiveness *(continued)*

and statistical significance of the effects of *Curiosity Corner* on children's performance.[9]

Oral language. Chambers et al. (2001) analyzed the differences between the *Curiosity Corner* and comparison groups for two measures (Mullen Scales of Early Learning [MSEL] Expressive Language Scale and MSEL Receptive Language Scale) in this outcome domain. The authors report findings separately for 3-year-old children in private child care centers and 4-year-old children in public school programs. The authors report a statistically significant effect of *Curiosity Corner* on Expressive Language for 3-year-old children but not for 4-year-old children. The effect size reported by Chambers et al. for the difference on the MSEL Expressive Language Scale between the 3-year-old *Curiosity Corner* and comparison groups is large enough to be considered substantively important according to WWC criteria (that is, at least 0.25). In WWC calculations, the differences are not statistically significant for either 3- or 4-year-old children on Expressive Language. For Receptive Language, Chambers et al. report, and the WWC confirms, the difference is not statistically significant or substantively important for 3- or 4-year-old children. The average effect size for oral language in this study, combining age groups and both measures in this domain, is not statistically significant or substantively important. PCER Consortium (2008) analyzed the effectiveness of *Curiosity Corner* on the Peabody Picture Vocabulary Test (PPVT-III) and the Test of Language Development (TOLD). The authors report, and the WWC confirms, differences between the treatment and control groups are not statistically significant or substantively important. According to WWC criteria, these two studies show no discernible effects on oral language.

Print knowledge. PCER Consortium (2008) analyzed *Curiosity Corner's* effectiveness on the Test of Early Reading Ability (TERA-3), the Woodcock-Johnson (WJ III) Letter Word Identification subtest, and the WJ III Spelling subtest. The study shows, and the WWC confirms, no statistically significant or substantively important differences between treatment and control groups on any of these measures, and thus, no discernible effects on print knowledge.

Phonological processing. PCER Consortium (2008) also examined *Curiosity Corner's* effects on the Elision subtest from the Preschool Comprehensive Test of Phonological Processing (Pre-CTOPPP). The study shows, and the WWC confirms, no statistically significant or substantively important differences between children who participated in *Curiosity Corner* and those in the control group, and thus, no discernible effects on phonological processing.

Cognition. Chambers et al. (2001) analyzed the differences between the *Curiosity Corner* and comparison groups for one measure (MSEL Visual Reception scale) in the cognition outcome domain. The authors report, and the WWC confirms, the difference between the intervention and comparison groups is not statistically significant for either 3- or 4-year-old children. According to WWC calculations, the effect sizes are not large enough to be considered substantively important and thus, according to WWC criteria, the study showed no discernible effects in this domain.

Math. PCER Consortium (2008) examined *Curiosity Corner's* effect on three math outcomes: the WJ III Applied Problems subtest, the Composite Score subtest from the Child Math Assessment-Abbreviated, and Shape Composition. The study shows, and the WWC confirms, none of the differences between children in the treatment and control groups are statistically significant or substantively important. Thus, according to WWC criteria, the study shows no discernible effects of *Curiosity Corner* on the math domain.

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, see the WWC Tutorial on Mismatch. For the formulas the WWC used to calculate the statistical significance, see Technical Details of WWC-Conducted Computations. For the *Curiosity Corner* study by Chambers et al. (2001) summarized here, a correction for clustering was needed. No corrections were needed for PCER Consortium (2008) because the analysis corrected for clustering by using HLM and no impacts were statistically significant.

Effectiveness *(continued)*

Rating of effectiveness

The WWC rates the effects of an intervention in a given outcome domain as positive, potentially positive, mixed, no discernible effects, potentially negative, or negative. The rating of effectiveness takes into account four factors: the quality of the research design, the statistical significance of the findings, the size of the difference between participants in the intervention and the comparison conditions, and the consistency in findings across studies (see the WWC Intervention Rating Scheme).

The WWC found *Curiosity Corner* to have no discernible effects on oral language, print knowledge, phonological processing, cognition, and math

Improvement index

The WWC computes an improvement index for each individual finding. In addition, within each outcome domain, the WWC computes an average improvement index for each study and an average improvement index across studies (see Technical Details of WWC-Conducted Computations). 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. Unlike the rating of effectiveness, the improvement index is based entirely on the size of the effect, regardless of the statistical significance of the effect, the study design, or the analyses. The improvement index can take on values between –50 and +50, with positive numbers denoting results favorable to the intervention group.

Curiosity Corner's average improvement index for oral language is +2 percentile points across the two studies, with a range of –3 to +14 percentile points. For print knowledge, one study showed an average improvement index of +3 percentile points, with a range of +2 to +4 percentile points across findings. For phonological processing, the average improvement index was +7 percentile points based on one study. For cognition, the average improvement index was –3 percentile points across findings in one study, with a range of –4 to –1 percentile points. The average improvement index for math was +4 percentile points from one study, with a range of 0 to +6 percentile points across findings in one study.

Summary

The WWC reviewed three studies of *Curiosity Corner.* One meets WWC evidence standards, and one meets WWC evidence standards with reservations. One study did not meet eligibility screens. Based on the two studies, the WWC found no discernible effects of *Curiosity Corner* on oral language, print knowledge, phonological processing, cognition, or math. The conclusions presented in this report may change as new research emerges.

References

Meets WWC evidence standards

Preschool Curriculum Evaluation Research (PCER) Consortium. (2008). Chapter 5. *Curiosity Corner*: Success for All Foundation. In *Effects of Preschool Curriculum Programs on School Readiness* (pp. 75–83). Washington, DC: National Center for Education Research, Institute of Education Sciences, U.S. Department of Education.

Meets WWC evidence standards with reservations

Chambers, B., Chamberlain, A., Hurley, E. A., & Slavin, R. E. (2001). *Curiosity Corner: Enhancing preschoolers' language abilities through comprehensive reform*. Paper presented at the Annual Meeting of the American Educational Research Association, Seattle, WA, April 2001.

Additional source:

Curiosity Corner summary of research: First year findings. Retrieved July 29, 2008, from http://www.successforall.com/_images/pdfs/410409000_CC_Research.pdf.

References *(continued)*

Studies that fall outside the Early Childhood Education protocol or do not meet WWC evidence standards

Chambers, B., Cheung, Alan C. K., & Slavin, R. E. (2006). Effective preschool programs for children at risk of school failure: A best evidence synthesis. In B. Spodek, & O. N. Saracho (Eds.), *Handbook of research on the education of young children* (Second ed., pp. 347-359). Mahwah, NJ: Lawrence Erlbaum Associates. The study is ineligible for review because it is not a primary analysis of the effectiveness of an intervention.

For more information about specific studies and WWC calculations, please see the WWC *Curiosity Corner* Technical Appendices.

Appendix

Appendix A1 Study Characteristics: Preschool Curriculum Evaluation Research Consortium, 2008 (randomized controlled trial)

Characteristic	Description
Study citation	Preschool Curriculum Evaluation Research (PCER) Consortium. (2008). Chapter 5. *Curiosity Corner*: Success for All Foundation. In *Effects of Preschool Curriculum Programs on School Readiness* (pp. 75–83). Washington, DC: National Center for Education Research, Institute of Education Sciences, U.S. Department of Education.
Participants	In this study, 18 preschools were randomly assigned to intervention (10 schools) or comparison (8 schools) conditions. Prior to random assignment, schools were sorted into blocks on a number of conditions, including teacher experience, school location, and state report card score. Random assignment occurred within each block. From the schools, 31 preschool classrooms participated in the study (14 intervention classrooms and 17 control classrooms). Participants included 215 preschool-age children whose parents consented to their participation in the study. At baseline, children were an average 4.7 years old, half were male, half were African-American, and 14% were reported as having a disability. Although the intervention and comparison groups were similar in race and disability status, the treatment group had more boys (61%) than the comparison group (38%), a difference that was statistically significant. Attrition from the analysis sample (children with parent consent) was low: 2% at baseline, 5% at end-of-preschool posttest, and 10% at end-of-kindergarten follow-up. Response rates varied by measure but were comparable across treatment and control groups.
Setting	The study was conducted in 18 schools (31 classrooms) in Florida, Kansas, and New Jersey.
Intervention	Intervention group children participated in *Curiosity Corner.* A Success For All (SFA) implementation measure was used by SFA trainers, who visited each classroom at least three times during the year and rated the implementation of each classroom. Fidelity of the classrooms was rated on a four point scale, ranging from "Not at all" (0) to "High" (3). The average fidelity score of the intervention classrooms was 2.0.
Comparison	The comparison condition varied across schools. Comparison schools in Florida primarily used the *Creative Curriculum*. The Kansas comparison schools participated in a blend of the *Preschool and Language Stimulation* curriculum and the *Animated Literacy* curriculum. Comparison schools in New Jersey used a teacher-developed curriculum. Comparison classrooms were visited twice a year by the trainers and rated using the same implementation measure as was used for the intervention classrooms. The average fidelity score of the comparison classrooms was 1.9.
Primary outcomes and measurement	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) Grammatic 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 Appendices A2.1–2.5.
Staff/teacher training	Success for All staff provided an initial training session for the intervention teachers and ongoing implementation support, including three visits a year to conduct observations and provide feedback.

Appendix A1.2 Study Characteristics: Chambers, Chamberlain, Hurley, and Slavin, 2001 (quasi-experimental design)

Characteristic	Description
Study citation	Chambers, B., Chamberlain, A., Hurley, E. A., & Slavin, R. E. (2001). *Curiosity Corner: Enhancing preschoolers' language abilities through comprehensive reform*. Paper presented at the Annual Meeting of the American Educational Research Association, Seattle, WA, April 2001.
Participants[1]	The study began with 448 low-income preschool children who ranged in age from two years, seven months to four years, eleven months. At posttest, 316 children were included in the study, with analysis samples ranging from 311 to 315. The three-year-olds were from private early childhood centers (n = 169), and the four-year-olds were from public preschools (n = 147). In the final sample, 68% of the children were African-American, 16% Caucasian, and 11% Hispanic; 49% were female. Eight preschools (public and private) were assigned to the *Curiosity Corner* intervention group, and eight preschools (public and private) matched on demographic characteristics were used as the comparison group.
Setting	The study took place in 16 preschools (a mix of public and private) in four high-poverty, urban school districts in New Jersey. All of the preschools were in Abbott districts and working to meet Abbott guidelines.
Intervention	The intervention group children participated in *Curiosity Corner* during the pilot year of the curriculum. *Curiosity Corner* was designed with 38 weekly thematic units. Additional information on duration, frequency, and intensity of implementation was not reported.
Comparison	The comparison group children participated in the regular early childhood curriculum at their preschool centers.
Primary outcomes and measurement	The primary outcome domains were children's oral language and cognition. The study used three subtests of a standardized test (the Mullen Scales of Early Learning, American Guidance Services Edition): expressive language, receptive language, and visual reception. The study also used the Early Childhood Environment Rating Scale-Revised (ECERS-R) to evaluate classroom quality, but the measure is not included in this WWC review because it is not relevant to the topic review. For a more detailed description of these outcome measures, see Appendices A2.1 and 2.4.
Staff/teacher training	The program provided teachers with detailed lesson instructions in the teacher's manual and materials for instructional activities. Teachers, teaching assistants, and administrators were trained in two-day initial training sessions, followed by six in-class visits by a Success for All Foundation (SFA) trainer. In addition, teachers were observed, mentored, and supported by *Curiosity Corner* coaches from the school districts, who were trained by SFA staff over a two-year period. Coaches also offered workshops to help teachers implement the curriculum.

1. Information on total sample size and the number of schools in each condition was provided by the study authors upon WWC request.

Appendix A2.1 Outcome measures for the oral language domain

Outcome measure	Description
Mullen Scales of Early Learning (MSEL) Expressive Language Scale	A scale from a standardized measure of children's expressive language skills, such as speaking and forming language (as cited in Chambers et al., 2001).
Mullen Scales of Early Learning Receptive Language Scale	A scale from a standardized measure of children's receptive language skills, such as auditory organization, sequencing, and use of spatial concepts (as cited in Chambers et al., 2001).
Peabody Picture Vocabulary Test-3rd Edition (PPVT-III)	A standardized measure of children's receptive vocabulary where children show 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 standardized measure 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).

Appendix A2.2 Outcome measures for the print knowledge domain

Outcome measure	Description
Test of Early Reading Ability III (TERA-3)	A standardized measure 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 standardized measure of identification of letters and reading of words (as cited in PCER Consortium, 2008).
Woodcock-Johnson III Spelling subtest	A standardized measure that assesses children's prewriting skills, such as drawing lines, tracing, and writing letters (as cited in PCER Consortium, 2008).

Appendix A2.3 Outcome measures for the phonological processing domain

Outcome measure	Description
Preschool Comprehensive Test of Phonological and Print Processing (Pre-CTOPPP), Elision subtest	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 (as cited in PCER Consortium, 2008).

Appendix A2.4 Outcome measures for the cognition domain

Outcome measure	Description
Mullen Scales of Early Learning Visual Reception Scale	A scale from a standardized measure of children's cognitive ability to process visual patterns (as cited in Chambers et al., 2001).

Appendix A2.5 Outcome measures for the math domain

Outcome measure	Description
Woodcock-Johnson III Applied Problems subtest	A standardized measure of children's ability to solve numerical and spatial problems, presented verbally with accompanying pictures of objects (as cited in PCER Consortium, 2008).
Child Math Assessment-Abbreviated (CMA-A) Composite Score	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 (as cited in PCER Consortium, 2008).
Building Blocks, Shape Composition task	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 (as cited in PCER Consortium, 2008).

Appendix A3.1 Summary of study findings included in the rating for the oral language domain[1]

Outcome measure	Study sample	Sample size (schools/ students)	Authors' findings from the study: Mean outcome[2] (standard deviation)[3] — *Curiosity Corner* group	Authors' findings from the study: Mean outcome[2] (standard deviation)[3] — Comparison group	WWC calculations: Mean difference[4] (*Curiosity Corner*—comparison)	WWC calculations: Effect size[5]	WWC calculations: Statistical significance[6] (at $\alpha = 0.05$)	WWC calculations: Improvement index[7]
Chambers et al., 2001 (quasi-experimental design)[8]								
MSEL Expressive Language Scale	3-year-olds	16/167[9]	39.26 (5.04)	37.54 (4.30)	1.72	0.36	ns	+14
MSEL Expressive Language Scale	4-year-olds	12/146[9]	43.58 (4.55)	43.29 (4.01)	0.29	0.07	ns	+3
MSEL Receptive Language Scale	3-year-olds	16/168[9]	37.76 (4.40)	37.52 (4.68)	0.24	0.05	ns	+2
MSEL Receptive Language Scale	4-year-olds	12/147[9]	43.10 (4.32)	42.85 (3.78)	0.25	0.06	ns	+2
Average for oral language (Chambers et al., 2001)[10]						**0.13**	**ns**	**+5**
PCER Consortium, 2008 (randomized controlled trial)[8]								
PPVT-III	Preschoolers	18/201	nr	nr	–0.17	–0.01	ns	0
TOLD-P:3 Grammatic Understanding subtest	Preschoolers	18/199	nr	nr	–0.38	–0.08	ns	–3
Average for oral language (PCER, 2008)[10]						**–0.05**	**ns**	**–2**
Domain average for oral language across all studies[10]						**0.04**	**na**	**+2**

ns = not statistically significant
na = not applicable
nr = not reported
MSEL = Mullen Scales of Early Learning
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 PCER Consortium (2008) are not included in these ratings but are reported in Appendix A4.1.
2. In the case of Chambers et al. (2001), posttest means are covariate-adjusted means. Chambers et al. (2001) included age and PPVT-III scores at pretest as covariates in the analysis.
3. 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.
4. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group. In the case of PCER Consortium (2008), the mean differences are covariate-adjusted.
5. 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.

Appendix A3.1 Summary of study findings included in the rating for the oral language domain[1] *(continued)*

6. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups.
7. 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.
8. 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 Chambers et al. (2001), a correction for clustering was needed, so the significance levels may differ from those reported in the original study. In the case of PCER Consortium (2008), no corrections were needed because the analysis corrected for clustering by using HLM and no impacts were statistically significant.
9. The sample size of schools was provided by the study authors at WWC request.
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.

Appendix A3.2 Summary of study findings included in the rating for the print knowledge domain[1]

Outcome measure	Study sample	Sample size (schools/ students)	Authors' findings from the study: Mean outcome (standard deviation)[2] — *Curiosity Corner* group	Authors' findings from the study: Mean outcome (standard deviation)[2] — Comparison group	WWC calculations: Mean difference[3] (*Curiosity Corner*—comparison)	WWC calculations: Effect size[4]	WWC calculations: Statistical significance[5] (at $\alpha = 0.05$)	WWC calculations: Improvement index[6]
PCER Consortium, 2008 (randomized controlled trial)[7]								
TERA-3	Preschoolers	18/200	nr	nr	0.83	0.10	ns	+4
WJ III Letter Word Identification subtest	Preschoolers	18/177	nr	nr	2.42	0.09	ns	+4
WJ III Spelling subtest	Preschoolers	18/194	nr	nr	0.97	0.04	ns	+2
Domain average for print knowledge[8]						**0.08**	**na**	**+3**

ns = not statistically significant
na = not applicable
nr = not reported
TERA-3 = Test of Early Reading Ability
WJ III = Woodcock-Johnson III

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 same 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.
3. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group. In the case of PCER Consortium (2008), the mean differences are covariate-adjusted.
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), no corrections were needed because the analysis corrected for clustering by using HLM and no impacts were statistically significant.
8. 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.

Appendix A3.3 Summary of study findings included in the rating for the phonological processing domain[1]

Outcome measure	Study sample	Sample size (schools/ students)	Authors' findings from the study: Mean outcome (standard deviation)[2] — *Curiosity Corner* group	Authors' findings from the study: Mean outcome (standard deviation)[2] — Comparison group	WWC calculations: Mean difference[3] (*Curiosity Corner*—comparison)	WWC calculations: Effect size[4]	WWC calculations: Statistical significance[5] (at $\alpha = 0.05$)	WWC calculations: Improvement index[6]
PCER Consortium, 2008 (randomized clinical trial)[7]								
Pre-CTOPPP Elision subtest	Preschoolers	18/204	nr	nr	–0.04	0.18	ns	+7
Domain average for phonological processing[8]						**0.18**	**na**	**+7**

ns = not statistically significant
na = not applicable
nr = not reported
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. Follow-up findings from the same 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. In the case of PCER Consortium (2008), the mean differences are covariate-adjusted.
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), no corrections for clustering or multiple comparisons were needed because the analysis corrected for clustering by using HLM and no impacts were statistically significant.
8. 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.

Appendix A3.4 Summary of study findings included in the rating for the cognition domain[1]

Outcome measure	Study sample	Sample size (schools/ students)	Authors' findings from the study: Mean outcome[2] (standard deviation)[3] — *Curiosity Corner* group	Authors' findings from the study: Mean outcome[2] (standard deviation)[3] — Comparison group	WWC calculations: Mean difference[4] (*Curiosity Corner*—comparison)	WWC calculations: Effect size[5]	WWC calculations: Statistical significance[6] (at $\alpha = 0.05$)	WWC calculations: Improvement index[7]
Chambers et al., 2001 (quasi-experimental design)[8]								
MSEL Visual Reception Scale	3-year-olds	16/165[9]	42.32 (3.54)	42.66 (4.04)	–0.34	–0.09	ns	–4
MSEL Visual Reception Scale	4-year-olds	12/146[9]	45.49 (3.20)	45.61 (3.20)	–0.12	–0.04	ns	–1
Domain average for cognition[10]						**–0.06**	**na**	**–3**

ns = not statistically significant
na = not applicable
MSEL = Mullen Scales of Early Learning

1. This appendix reports findings considered for the effectiveness rating and the average improvement indices for the cognition domain.
2. In the case of Chambers et al. (2001), posttest means are covariate-adjusted means. Chambers et al. (2001) included age and PPVT-III scores at pretest as covariates in the analysis.
3. 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.
4. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group.
5. For an explanation of the effect size calculation, see Technical Details of WWC-Conducted Computations.
6. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups.
7. 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.
8. 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 Chambers et al. (2001), a correction for clustering was needed, so the significance levels may differ from those reported in the original study.
9. The number of schools was provided by study authors at WWC request. The sample size for the comparison group of 4-year-olds reported in the original study was incorrect and the correct sample size was provided by the study authors.
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.

Appendix A3.5 Summary of study findings included in the rating for the math domain[1]

Outcome measure	Study sample	Sample size (schools/ students)	Authors' findings from the study: Mean outcome (standard deviation)[2] — *Curiosity Corner* group	Authors' findings from the study: Mean outcome (standard deviation)[2] — Comparison group	WWC calculations: Mean difference[3] (*Curiosity Corner*—comparison)	WWC calculations: Effect size[4]	WWC calculations: Statistical significance[5] (at $\alpha = 0.05$)	WWC calculations: Improvement index[6]
PCER Consortium, 2008 (randomized controlled trial)[7]								
WJ III Applied Problems	Preschoolers	18/180	nr	nr	1.90	0.10	ns	+4
CMA-A Composite	Preschoolers	18/204	nr	nr	0.00	0.01	ns	0
Shape Composition	Preschoolers	18/200	nr	nr	0.15	0.16	ns	+6
Domain average for math[8]						**0.09**	**na**	**+4**

ns = not statistically significant
na = not applicable
nr = not reported
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 same study are not included in these ratings, but are reported in Appendix A4.4.
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. In the case of PCER Consortium (2008), the mean differences are covariate-adjusted.
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 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), no corrections were needed because the analysis corrected for clustering by using HLM and no impacts were statistically significant.
8. 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.

Appendix A4.1 Summary of kindergarten follow-up findings for the oral language domain[1]

Outcome measure	Study sample	Sample size (schools/ students)	Authors' findings from the study: Mean outcome (standard deviation)[2] — *Curiosity Corner* group	Authors' findings from the study: Mean outcome (standard deviation)[2] — Comparison group	WWC calculations: Mean difference[3] (*Curiosity Corner*—comparison)	WWC calculations: Effect size[4]	WWC calculations: Statistical significance[5] (at $\alpha = 0.05$)	WWC calculations: Improvement index[6]
PCER Consortium, 2008 (randomized controlled trial)[7]								
PPVT-III	Kindergarteners	69/189	nr	nr	2.42	0.14	ns	+6
TOLD-P:3 Grammatic Understanding subtest	Kindergarteners	69/190	nr	nr	0.74	0.15	ns	+6

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 for measures that fall in the oral language domain. Posttest scores for preschoolers 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. In the case of PCER Consortium (2008), the mean differences are covariate-adjusted.
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 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. The WWC does not make corrections for multiple comparisons for follow-up findings.

Appendix A4.2 Summary of kindergarten follow-up findings for the print knowledge domain[1]

Outcome measure	Study sample	Sample size (schools/ students)	Authors' findings from the study: Mean outcome (standard deviation)[2] — *Curiosity Corner* group	Authors' findings from the study: Mean outcome (standard deviation)[2] — Comparison group	WWC calculations: Mean difference[3] (*Curiosity Corner*—comparison)	WWC calculations: Effect size[4]	WWC calculations: Statistical significance[5] (at $\alpha = 0.05$)	WWC calculations: Improvement index[6]
PCER Consortium, 2008 (randomized controlled trial; kindergarten follow-up)[7]								
TERA-3	Kindergarteners	69/188	nr	nr	3.50	0.43	Statistically significant	+17
WJ III Letter Word Identification subtest	Kindergarteners	69/189	nr	nr	11.26	0.43	Statistically significant	+17
WJ III Spelling subtest	Kindergarteners	69/182	nr	nr	5.60	0.20	ns	+8

ns = not statistically significant
nr = not reported
TERA-3 = Test of Early Reading Ability
WJ III = Woodcock-Johnson III

1. This appendix presents follow-up findings for measures that fall in the print knowledge domain. Posttest scores for preschoolers were used for rating purposes and are 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. In the case of PCER Consortium (2008), the mean differences are covariate-adjusted.
4. For an explanation of the effect size calculation, see Technical Details of WWC-Conducted Computations.
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. The WWC does not make corrections for multiple comparisons for follow-up findings.

Appendix A4.3 Summary of kindergarten follow-up findings for the phonological processing domain[1]

Outcome measure	Study sample	Sample size (schools/ students)	Authors' findings from the study: Mean outcome (standard deviation)[2] — *Curiosity Corner* group	Authors' findings from the study: Mean outcome (standard deviation)[2] — Comparison group	WWC calculations: Mean difference[3] (*Curiosity Corner*—comparison)	WWC calculations: Effect size[4]	WWC calculations: Statistical significance[5] (at $\alpha = 0.05$)	WWC calculations: Improvement index[6]
PCER Consortium, 2008 (randomized controlled trial; kindergarten follow-up)[7]								
CTOPP Elision subtest	Kindergarteners	69/193	nr	nr	0.94	0.25	ns	+10

ns = not statistically significant
nr = not reported
CTOPP = Comprehensive Test of Phonological Processing

1. This appendix presents follow-up findings for measures that fall in the phonological processing domain. Posttest scores for preschoolers were used for rating purposes and are presented in Appendix A3.3.
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. In the case of PCER Consortium (2008), the mean differences are covariate-adjusted.
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. The effect size for the CTOPP measure was calculated using an ANCOVA model and is not comparable to the effect sizes for other outcomes in PCER, which were calculated with repeated measures models.
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. The WWC does not make corrections for multiple comparisons for follow-up findings.

Appendix A4.4 Summary of kindergarten follow-up findings for the math domain[1]

Outcome measure	Study sample	Sample size (schools/ students)	Authors' findings from the study: Mean outcome (standard deviation)[2] — *Curiosity Corner* group	Authors' findings from the study: Mean outcome (standard deviation)[2] — Comparison group	WWC calculations: Mean difference[3] (*Curiosity Corner*—comparison)	WWC calculations: Effect size[4]	WWC calculations: Statistical significance[5] (at $\alpha = 0.05$)	WWC calculations: Improvement index[6]
PCER Consortium, 2008 (randomized controlled trial; kindergarten follow-up)[7]								
WJ III Applied Problems	Kindergarteners	69/188	nr	nr	5.11	0.26	ns	+10
CMA-A Composite	Kindergarteners	69/194	nr	nr	–0.01	–0.05	ns	–2
Shape Composition	Kindergarteners	69/194	nr	nr	0.31	0.32	ns	+13

ns = not statistically significant
nr = not reported
WJ III = Woodcock-Johnson III
CMA-A = Child Math Assessment-Abbreviated

1. This appendix presents follow-up findings for measures that fall in the math domain. Posttest scores for preschoolers 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. In the case of PCER Consortium (2008), the mean differences are covariate-adjusted.
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. The WWC does not make corrections for multiple comparisons for follow-up findings.

Appendix A5.1 *Curiosity Corner* rating for the oral language domain

The WWC rates an intervention's effects for a given outcome domain as positive, potentially positive, mixed, no discernible effects, potentially negative, or negative.[1]

For the outcome domain of oral language, the WWC rated *Curiosity Corner* as having no discernible effects. The remaining ratings (potentially negative effects and negative effects) were not considered, as *Curiosity Corner* 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. Two studies of *Curiosity Corner* showed no statistically significant or substantively important effects, either positive or negative.

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. Two studies of *Curiosity Corner* showed no statistically significant or substantively important positive effects.

AND

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

  Met. Neither of the two studies of *Curiosity Corner* showed a statistically significant or substantively important negative effect.

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. Two studies of *Curiosity Corner* showed no 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.

  Not met. Two studies of *Curiosity Corner* showed no 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. Two studies of *Curiosity Corner* showed no 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. Two studies of *Curiosity Corner* showed no 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 *Curiosity Corner* rating for the print knowledge domain

The WWC rates an intervention's effects for a given outcome domain as positive, potentially positive, mixed, no discernible effects, potentially negative, or negative.[1]

For the outcome domain of print knowledge, the WWC rated *Curiosity Corner* as having no discernible effects. The remaining ratings (potentially negative effects and negative effects) were not considered, as *Curiosity Corner* 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. One study of *Curiosity Corner* showed no statistically significant or substantively important effects, either positive or negative.

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. One study of *Curiosity Corner* showed no statistically significant or substantively important positive effects.

AND

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

  Met. No study of *Curiosity Corner* showed a statistically significant or substantively important negative effect.

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. One study of *Curiosity Corner* showed no 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.

  Not met. One study of *Curiosity Corner* showed no 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. One study of *Curiosity Corner* showed no 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. One study of *Curiosity Corner* showed no 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 *Curiosity Corner* rating for the phonological processing domain

The WWC rates an intervention's effects for a given outcome domain as positive, potentially positive, mixed, no discernible effects, potentially negative, or negative.[1]

For the outcome domain of phonological processing, the WWC rated *Curiosity Corner* as having no discernible effects. The remaining ratings (potentially negative effects and negative effects) were not considered, as *Curiosity Corner* 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. One study of *Curiosity Corner* showed no statistically significant or substantively important effects, either positive or negative.

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. One study of *Curiosity Corner* showed no statistically significant or substantively important positive effects.

AND

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

  Met. No study of *Curiosity Corner* showed a statistically significant or substantively important negative effect.

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. One study of *Curiosity Corner* showed no 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.

  Not met. One study of *Curiosity Corner* showed no 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. One study of *Curiosity Corner* showed no 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. One study of Curiosity Corner showed no 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 *Curiosity Corner* rating for the cognition domain

The WWC rates an intervention's effects for a given outcome domain as positive, potentially positive, mixed, no discernible effects, potentially negative, or negative.[1]

For the outcome domain of cognition, the WWC rated *Curiosity Corner* as having no discernible effects. The remaining ratings (potentially negative effects and negative effects) were not considered, as *Curiosity Corner* 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. One study of *Curiosity Corner* showed no statistically significant or substantively important effects, either positive or negative.

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. One study of *Curiosity Corner* showed no statistically significant or substantively important positive effects.

AND

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

  Met. No study of *Curiosity Corner* showed a statistically significant or substantively important negative effect.

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. One study of *Curiosity Corner* showed no 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.

  Not met. One study of *Curiosity Corner* showed no 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. One study of *Curiosity Corner* showed no 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. One study of *Curiosity Corner* showed no 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 *Curiosity Corner* rating for the math domain

The WWC rates an intervention's effects for a given outcome domain as positive, potentially positive, mixed, no discernible effects, potentially negative, or negative.[1]

For the outcome domain of math, the WWC rated *Curiosity Corner* as having no discernible effects. The remaining ratings (potentially negative effects and negative effects) were not considered, as *Curiosity Corner* 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. No study of *Curiosity Corner* showed statistically significant or substantively important effects, either positive or negative.

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. One study of *Curiosity Corner* showed no statistically significant or substantively important positive effects.

AND

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

  Met. No study of *Curiosity Corner* showed a statistically significant or substantively important negative effect.

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. One study of *Curiosity Corner* showed no 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.

  Not met. One study of *Curiosity Corner* showed no 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. One study of *Curiosity Corner* showed no 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. One study of *Curiosity Corner* showed no 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[1]
		Schools	Students	
Oral language[2]	2	34	527	Medium to large
Print knowledge[2]	1	18	211	Small
Phonological processing	1	18	211	Small
Cognition	1	16	316	Small
Math[2]	1	18	211	Small
Early reading and writing	0	na	na	na

na = not applicable

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."
2. Sample size varies by outcome measure.