What Works Clearinghouse™

Ladders to Literacy

Program Description[1]

Ladders to Literacy is a supplemental early literacy curriculum composed of 60 activities designed to develop children's print/book awareness, metalinguistic awareness, and oral language skills. The *Ladders to Literacy* activities can be implemented in a variety of early childhood settings and adapted for children with special needs. Although a *Ladders to Literacy* curriculum is also available for kindergarten students, this intervention report focuses on the preschool *Ladders to Literacy* supplemental early literacy curriculum.

Research[2]

The What Works Clearinghouse (WWC) identified two studies of *Ladders to Literacy* that both fall within the scope of the Early Childhood Education topic area and meet WWC evidence standards.[3] One study meets WWC evidence standards without reservations and one study meets WWC evidence standards with reservations, and together, they included 139 children in 26 preschool classrooms in southern New Hampshire.

The WWC considers the extent of evidence for *Ladders to Literacy* on the school readiness of preschool children to be small for four outcome domains—oral language, print knowledge, phonological processing, and math. There were no studies that meet standards in early reading and writing, and cognition, so we do not report on the effectiveness of *Ladders to Literacy* for those domains in this intervention report.

Report Contents

Effectiveness

Ladders to Literacy was found to have potentially negative effects on oral language and no discernible effects on print knowledge, phonological processing, and math for preschool children.

Table 1. Summary of findings[4]

Outcome domain	Rating of effectiveness	Improvement index *(percentile points)*		Number of studies	Number of children	Extent of evidence
		Average	Range			
Oral language	Potentially negative effects	–7	–15 to +2	2	139	Small
Print knowledge	No discernible effects	0	–12 to +11	1	105	Small
Phonological processing	No discernible effects	–6	na	1	105	Small
Math	No discernible effects	+3	+1 to +4	1	105	Small

na = not applicable

Program Information

Background

Ladders to Literacy was developed by Angela Notari-Syverson, Rollanda E. O'Connor, and Patricia F. Vadasy, and is distributed by Brookes Publishing Company. Address: P.O. Box 10624, Baltimore, MD 21285-0624. Email: custserv@brookespublishing.com. Web: http://www.brookespublishing.com. Telephone: (800) 638-3775.

Program details

Ladders to Literacy is a supplemental early childhood curriculum that focuses on developing children's early language and literacy skills. The supplemental curriculum addresses three components of literacy development: print/book awareness, metalinguistic awareness, and oral language skills. The print/book awareness component includes activities such as drawing, pretending to write, and creating graphic representations. The metalinguistic awareness component focuses on the identification of sounds, phonemes, and rhymes through lessons such as "Clap the Syllables" and "First Sound Song." The oral language component includes activities designed to enhance vocabulary development and to engage children in conversations in which they respond to open-ended questions and reconstruct past experiences. There are 20 lessons for each of the three *Ladders to Literacy* components.

The activities that are to be conducted in each lesson, the rationale for the activities, overall goals, targeted skills, and required materials, are described in a teacher's guide, *Ladders to Literacy: A Preschool Activity Book, Second Edition*. Teachers prepare the *Ladders to Literacy* activities and assemble the required materials according to the specifications from this guide. The guide describes how to individualize activities for children with varying skill levels and how to adapt lessons for children with special needs. The guide also includes an overview of the theoretical framework underlying *Ladders to Literacy* and a literacy checklist to help teachers monitor children's progress. In addition to the teacher-led activities, an appendix provides a variety of activities that parents and children can do together at home to reinforce skills being taught in the classroom.

Cost

The cost of the teacher's guide, *Ladders to Literacy: A Preschool Activity Book, Second Edition*, is $49.95. Only one guide is necessary per teacher or classroom. There are no additional materials or activity books for children to use in the classroom.

Research Summary

The WWC identified 13 studies that investigated the effects of *Ladders to Literacy* on the school readiness of preschool children.

Table 2. Scope of reviewed research

Grade	PK
Delivery method	Whole class
Program type	Supplement

The WWC reviewed four of those studies against group design evidence standards. One study (Russell, 2005) is a randomized controlled trial that meets WWC evidence standards without reservations and one study (Preschool Curriculum Evaluation Research [PCER] Consortium, 2008) is a randomized controlled trial that meets WWC evidence standards with reservations. Those two studies are summarized in this report.[5] Two studies do not meet WWC evidence standards. The remaining nine studies do not meet WWC eligibility screens for review in this topic area. Citations for all 13 studies are in the References section, which begins on p. 8.

Summary of study meeting WWC evidence standards without reservations

Russell (2005) conducted a randomized controlled trial of 4-year-old children from 12 Head Start classrooms in southern New Hampshire. The classrooms in this study were recruited for the pilot year of the national PCER Consortium (2008) study.[6] The 12 classrooms in Russell (2005) were randomly assigned to either an intervention group, which implemented *The Creative Curriculum*® supplemented with *Ladders to Literacy*, or to a comparison group, which implemented *The Creative Curriculum*® by itself. Thirty-four children participated in the study (18 in the intervention group and 16 in the comparison group). At baseline, the children in the study averaged 4.7 years of age, and none were identified as having a disability. Because of difficulties in obtaining parental consent for study participation, Russell (2005) used a posttest-only design to investigate effects on outcomes in the oral language domain during the 2002–03 school year (the pilot year of the national PCER Consortium [2008] study). The children were assessed after at least four months of exposure to the curriculum.

Summary of study meeting WWC evidence standards with reservations

The PCER Consortium (2008) study assessed the effects of *Ladders to Literacy* using a randomized controlled trial of 14 Head Start preschool classrooms in southern New Hampshire. In the pilot year of the study the classrooms were randomly assigned to either an intervention group, which implemented *Ladders to Literacy* as a supplement to *The Creative Curriculum*®, or to a comparison group, which implemented *The Creative Curriculum*® without the *Ladders to Literacy* supplement. Eleven of the 14 classrooms were randomly assigned in the pilot year of the PCER Consortium (2008) study and three other classrooms (in the same centers) were added to the sample for the national PCER evaluation year.

Although the PCER Consortium (2008) study used a randomized controlled trial design to assign classrooms to intervention or comparison conditions in the pilot study year, the study analyzed data from the national PCER evaluation year (2003–04 school year), when children who had been in the classrooms at the time of random assignment (the start of the 2002–03 school year) had moved on to kindergarten, and a new class of children had replaced them. Thus, the study has high attrition at the child level and, under WWC standards, must demonstrate baseline equivalence between the intervention and comparison group sample of children used in the analyses of outcomes.

The authors investigated effects on oral language, print knowledge, phonological processing, and math. The outcome measures examined in this study were not examined in the Russell (2005) study described above. The WWC based its effectiveness ratings on findings from comparisons of 54 children who received *Ladders to Literacy* and 51 comparison group children. Children in the sample were 4.6 years old on average, with 44% male, 25% reported as having a disability, 39% Caucasian, 31% Hispanic, and 11% African American. The study demonstrated the baseline equivalence of the outcome measures in the oral language, print knowledge, phonological processing,

and math domains for the analytic sample of intervention and comparison group children at the end of the preschool year.[7] The authors reported on the effects of *Ladders to Literacy* in the spring of the preschool year and again at the end of kindergarten. The kindergarten findings are not reported here because information about the baseline equivalence of the outcome measures for the kindergarten sample was not provided in the report. The authors also reported findings on the Social Skills Rating Scale; however, these findings are not reported here because the current Early Childhood Education topic area protocol does not include sociobehavioral outcomes.

Effectiveness Summary

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

Summary of effectiveness for the oral language domain

One study that met WWC standards without reservations and one study that met WWC standards with reservations reported findings in the oral language domain.

Russell (2005) analyzed the effectiveness of *Ladders to Literacy* on oral language using two measures: mean length of utterance (MLU) and type token ratio (TTR). The author did not find statistically significant differences between the *Ladders to Literacy* group and the comparison group on either measure, and the effect sizes were not large enough to be considered substantively important according to WWC criteria (that is, an effect size of at least 0.25). The WWC characterizes these study findings as an indeterminate effect.

The PCER Consortium (2008) analyzed the effectiveness of *Ladders to Literacy* on oral language using the Peabody Picture Vocabulary Test–III (PPVT-III) and the Test of Oral Language Development–Primary III (TOLD-P:3) Grammatic Understanding subtest. The authors reported that differences between the *Ladders to Literacy* group and the comparison group were not statistically significant on either of these measures. The effect size for the TOLD-P:3 was not large enough to be considered substantively important according to WWC criteria. However, there was a substantively important negative effect of –0.38 on the PPVT-III, and the mean effect for the oral language domain, –0.30, was also substantively important. The WWC characterizes these study findings as a substantively important negative effect.

Thus, for the oral language domain, one study found an indeterminate effect and one study found a substantively important negative effect. This results in a rating of potentially negative effects, with a small extent of evidence.

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

Rating of effectiveness	Criteria met
Potentially negative effects *Evidence of a negative effect with no overriding contrary evidence.*	In the two studies that reported findings, the estimated impact of the intervention on outcomes in the *oral language* domain was an indeterminate effect in one study and a substantively important negative effect in another study.
Extent of evidence	**Criteria met**
Small	Two studies that included 139 children in eight schools reported evidence of effectiveness in the *oral language* domain.

Summary of effectiveness for the print knowledge domain

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

The PCER Consortium (2008) analyzed the effectiveness of *Ladders to Literacy* on print knowledge using the Test of Early Reading Ability–III (TERA-3), the Woodcock-Johnson III (WJ-III) Letter-Word Identification subtest, and the WJ-III Spelling subtest. The authors reported that differences between the *Ladders to Literacy* group and the comparison group were not statistically significant on any of these measures. The effect size for the WJ-III Letter-Word Identification subtest was not large enough to be considered substantively important according to WWC criteria. However, there was a substantively important negative effect of –0.30 on the TERA-3 and a substantively important positive effect of 0.27 on the WJ-III Spelling subtest. The mean effect for the print knowledge domain was neither statistically significant nor substantively important. The WWC characterizes these study findings as an indeterminate effect.

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

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

Rating of effectiveness	Criteria met
No discernible effects *No affirmative evidence of effects.*	In the one study that reported findings, the estimated impact of the intervention on outcomes in the *print knowledge* domain was neither statistically significant nor large enough to be substantively important.
Extent of evidence	**Criteria met**
Small	One study that included 105 children in eight schools reported evidence of effectiveness in the *print knowledge* domain.

Summary of effectiveness for the phonological processing domain

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

The PCER Consortium (2008) analyzed the effectiveness of *Ladders to Literacy* on phonological processing using the Preschool Comprehensive Test of Phonological and Print Processing (Pre-CTOPPP) Elision subtest. The authors reported that the difference between the *Ladders to Literacy* group and the comparison group was not statistically significant and not large enough to be substantively important according to WWC criteria. The WWC characterizes these study findings as an indeterminate effect.

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

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

Rating of effectiveness	Criteria met
No discernible effects *No affirmative evidence of effects.*	In the one study that reported findings, the estimated impact of the intervention on outcomes in the *phonological processing* domain was neither statistically significant nor large enough to be substantively important.
Extent of evidence	**Criteria met**
Small	One study that included 105 children in eight schools reported evidence of effectiveness in the *phonological processing* domain.

Summary of effectiveness for the math domain

One study that meets WWC evidence standards with reservations reported findings in the math domain.

The PCER Consortium (2008) analyzed the effectiveness of *Ladders to Literacy* on math using the WJ-III Applied Problems subtest, the Child Math Assessment–Abbreviated (CMA-A), and the Building Blocks Shape Composition task. The authors reported that differences between the *Ladders to Literacy* group and the comparison group were not statistically significant and not large enough to be considered substantively important according to WWC criteria. The WWC characterizes these study findings as an indeterminate effect.

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

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

Rating of effectiveness	Criteria met
No discernible effects *No affirmative evidence of effects.*	In the one study that reported findings, the estimated impact of the intervention on outcomes in the *math* domain was neither statistically significant nor large enough to be substantively important.
Extent of evidence	**Criteria met**
Small	One study that included 105 children in eight schools reported evidence of effectiveness in the *math* domain.

References

Study that meets WWC evidence standards without reservations

Russell, J. (2005). *An investigation of preschool oral language improvements through Ladders to Literacy* (Unpublished master's thesis). University of New Hampshire, Durham. (62329791)

Study that meets WWC evidence standards with reservations

Preschool Curriculum Evaluation Research (PCER) Consortium. (2008). Creative Curriculum with Ladders to Literacy: University of New Hampshire. In *Effects of preschool curriculum programs on school readiness* (pp. 65–73). Washington, DC: National Center for Education Research, Institute of Education Sciences, U.S. Department of Education.

Studies that do not meet WWC evidence standards

Good, J. L. (2003). Developing early literacy skills in young children with symptoms of inattention and hyperactivity. *Dissertation Abstracts International, 64*(06A), 106-1966. The study does not meet WWC evidence standards because the measures of effectiveness cannot be attributed solely to the intervention—there was only one unit assigned to one or both conditions.

Notari-Syverson, A. (1999). *Supporting early literacy development in young children with disabilities: A comprehensive interactive emergent literacy curriculum for preschoolers* (Final report to the U.S. Department of Education). Seattle, WA: Washington Research Institute. The study does not meet WWC evidence standards because it uses a quasi-experimental design in which the analytic intervention and comparison groups are not shown to be equivalent.

Additional source:

Notari-Syverson, A., O'Connor, R. E., & Vadasy, P. F. (1996, April). *Facilitating language and literacy development in preschool children: To each according to their needs*. Paper presented at the American Educational Research Association Meeting, New York. (ERIC Document ED395692)

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

Barker, R. M., Saunders, K. J., & Brady, N. C. (2012). Reading instruction for children who use AAC: Considerations in the pursuit of generalizable results. *AAC: Augmentative and Alternative Communication, 28*(3), 160–170. The study is ineligible for review because it is a secondary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.

Chambers, B., Cheung, A., Slavin, R. E., Smith, D., & Laurenzano, M. (2010). *Effective early childhood education programs: A systematic review*. Baltimore, MD: Johns Hopkins University, Center for Research and Reform in Education. Retrieved from http://www.bestevidence.org/word/early_child_ed_Sep_22_2010.pdf The study is ineligible for review because it is a secondary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.

Coston, J. H. (1997). *The effects of a comprehensive curriculum on literacy development* (Unpublished master's thesis). Valdosta State University, GA. The study is ineligible for review because it does not examine an intervention implemented in a way that falls within the scope of the review.

Ehri, L. C., Nunes, S. R., Willows, D. M., Schuster, B. V., Yaghoub-Zadeh, Z., & Shanahan, T. (2001). Phonemic awareness instruction helps children learn to read: Evidence from the national reading panel's meta-analysis. *Reading Research Quarterly, 36*(3), 250–287. The study is ineligible for review because it does not examine an intervention implemented in a way that falls within the scope of the review.

McKnight, C., Lee, S., & Schowengerdt, R. (2001). *Effects of specific strategy training on phonemic awareness and reading aloud with preschoolers: A comparison study*. Retrieved from ERIC database. (ED452518) The study

is ineligible for review because it does not examine an intervention implemented in a way that falls within the scope of the review.

Notari-Syverson, A. (2005). *Ladders to Literacy: Outreach project. Final grant performance report* (Submitted to the U.S. Department of Education, No. H324R000008). Seattle, WA: Washington Research Institute. The study is ineligible for review because it does not use a comparison group design or a single-case design.

Notari-Syverson, A. (2007). *Model demonstration project for children with disabilities: Final grant performance report* (Submitted to the U.S. Department of Education, No. H324M020084). Seattle, WA: Washington Research Institute. The study is ineligible for review because it does not examine an intervention implemented in a way that falls within the scope of the review.

Slocum, T. A. (1993). Teaching phonological awareness to young children with learning disabilities. *Exceptional Children, 59*(6), 532–546. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample includes less than 50% general education students.

Vaughn, S., Linan-Thompson, S., & Hickman, P. (2003). Response to instruction as a means of identifying students with reading/learning disabilities. *Exceptional Children, 69*(4), 391–409. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample includes less than 50% general education students.

Appendix A.1: Research details for Russell, 2005

Russell, J. (2005). *An investigation of preschool oral language improvements through Ladders to Literacy* (Unpublished master's thesis). University of New Hampshire, Durham. (62329791)

Table A1. Summary of findings

Meets WWC evidence standards without reservations

		Study findings	
Outcome domain	Sample size	Average improvement index *(percentile points)*	Statistically significant
Oral language	12 classrooms/34 children	–1	No

Setting

The study took place in 12 Head Start classrooms in two urban and four rural areas in southern New Hampshire.

Study sample

This study was a posttest-only design (no pretest was possible due to difficulties in obtaining parental consent for study participation). The study was conducted with children from 12 Head Start classrooms in the 2002–03 school year. Twelve teachers participated in the study. The classrooms were selected in 2002 from a list of prospective study participants and then randomly assigned to either an intervention or a comparison group. The researchers first identified four urban full-day classrooms and randomly assigned two to the intervention group and two to the comparison group. They also selected (a) two urban half-day classrooms with high numbers of Spanish-speaking children, (b) two additional urban half-day classrooms, (c) two suburban/rural classrooms from towns with a kindergarten program, and (d) two classrooms from towns with no kindergarten program. From each group, one classroom was randomly assigned to the intervention and one to the comparison group. Study eligibility was limited to children speaking English as their primary language and not enrolled in a special education program. Among children meeting these eligibility criteria, the study author randomly selected 60 children to participate (33 intervention and 27 comparison). Of the 60 children selected for the study, 34 children received parental consent to participate in the study (18 intervention and 16 comparison). The analysis sample meets attrition standards for the Early Childhood Education topic area, as described in its review protocol. At study enrollment, children in the analysis sample averaged 4.7 years of age, 65% were male, 71% were Caucasian, 12% were Hispanic, 6% were African American, and none of the children were identified as having a disability.

Intervention group

Intervention classrooms implemented *Ladders to Literacy* as a supplementary curriculum to *The Creative Curriculum*®. Teachers were trained to implement 18 language and literacy activities (of 60 that were available) across three domains (print/book awareness, metalinguistic awareness, and oral language). Fidelity of implementation was assessed twice during the study year, first in January/February 2003 and again in March/April 2003. For both the intervention (*Ladders to Literacy* plus *The Creative Curriculum*®) and comparison classrooms (*The Creative Curriculum*® alone), fidelity for *The Creative Curriculum*® was assessed using a checklist published by *The Creative Curriculum*® publishers. Fidelity to the *Ladders to Literacy* curriculum was assessed using an implementation checklist prepared by the *Granite Ladders* project staff. In the intervention group, implementation of both curricula was characterized as "moderate," averaging 52%–61% of *The Creative Curriculum*® activities and 41%–54% of the *Ladders to Literacy* activities.

Comparison group

The comparison group implemented *The Creative Curriculum*® without *Ladders to Literacy*. *The Creative Curriculum*® is a comprehensive curriculum for 3- to 5-year-old children. It addresses four areas of development: social/emotional, physical, cognitive, and language development. The curriculum required the physical space of the classroom to be structured into 10 interest areas: blocks, dramatic play, toys and games, art, library, discovery, sand and water, music and movement, cooking, and computers. Time was also allotted for outdoor activities. The 10 interest areas were designed to address curriculum content such as literacy, math, science, social studies, the arts, and technology, as well as process skills such as observing, exploring, and problem solving. Fidelity of implementation of *The Creative Curriculum*® in the comparison group classrooms was assessed using a checklist published by *The Creative Curriculum*® publishers. Implementation was characterized as "moderate" for comparison group classrooms, with teachers implementing 46%–48% of the strategies included in *The Creative Curriculum*®.

Outcomes and measurement

To measure oral language for posttests, researchers analyzed samples of children's speech and the MLU and TTR calculations. The children were assessed after at least four months of exposure to the curriculum. For a more detailed description of these outcome measures, see Appendix B.

Support for implementation

Both intervention and comparison teachers received at least one day of training on *The Creative Curriculum*®. Intervention group teachers received an additional 2 days of training on *Ladders to Literacy* activities in early fall 2002. Technical assistance to implement *Ladders to Literacy* activities was available to the intervention teachers, if needed.

Appendix A.2: Research details for PCER Consortium, 2008

Preschool Curriculum Evaluation Research (PCER) Consortium. (2008). Creative Curriculum with Ladders to Literacy: University of New Hampshire. In *Effects of preschool curriculum programs on school readiness* (pp. 65–73). Washington, DC: National Center for Education Research, Institute of Education Sciences, U.S. Department of Education.

Table A2. Summary of findings

Meets WWC evidence standards with reservations

		Study findings	
Outcome domain	Sample size	Average improvement index *(percentile points)*	Statistically significant
Oral language	14 classrooms/105 children	–12	No
Print knowledge	14 classrooms/105 children	0	No
Phonological processing	14 classrooms/105 children	–6	No
Math	14 classrooms/105 children	+3	No

Setting

The study was conducted in 14 Head Start classrooms in New Hampshire.

Study sample

This randomized controlled study, conducted during the 2003–04 and 2004–05 school years, included an intervention group that implemented *The Creative Curriculum*® supplemented by the *Ladders to Literacy* curriculum and a comparison group that implemented just *The Creative Curriculum*®. In 2002–03 (the study's pilot year) the researchers recruited 12 Head Start classrooms to participate in the study, grouped them according to whether their respective Head Start programs are located in a rural or urban area and are full- or half-day programs, and then randomly assigned the classrooms within each group to intervention and comparison groups. In the study's evaluation year (2003–04), 11 of the pilot-year classrooms and nine of the teachers were retained. One comparison classroom was replaced with another classroom from the same center, and two additional classrooms were randomly assigned to the intervention and comparison groups. This resulted in a sample of 14 classrooms (seven intervention and seven comparison). For most of the classrooms, the intervention condition had been in place for a full year when the evaluation year started. After parental consent was obtained, the sample included 123 children at baseline; 105 children were included in the final sample (54 intervention and 51 comparison). Baseline equivalence between the analytic sample of intervention and comparison children was established from data on baseline outcome measures provided by the study authors. At baseline, children in the study averaged 4.6 years of age, 44% were male, 39% were Caucasian, 31% were Hispanic, and 11% were African American.

Intervention group

Intervention classrooms implemented *Ladders to Literacy* as a supplementary curriculum to *The Creative Curriculum*®. Researchers selected 27 of the 60 *Ladders to Literacy* activities for implementation in the classrooms assigned to the intervention group. Teachers were trained to implement those 27 language and literacy activities across three domains (print/book awareness, metalinguistic awareness, and oral language). In November and December 2003, teachers were expected to implement nine activities (three from each of the three major domains). For the rest of the preschool year (January to May 2004), teachers were expected to continue implementing those nine activities and to implement an additional three to six activities each

month so that, by May 2004, teachers had implemented all 27 activities. Fidelity of implementation was assessed by conducting observations from December 2003 through April 2004 in the classrooms assigned to use the *Ladders to Literacy* curriculum. Researchers used a global fidelity measure to rate the overall fidelity with which the curriculum was implemented. On a four-point scale (0 = "not at all" to 3 = "high"), classrooms implementing the *Ladders to Literacy* curriculum were rated in the high-medium range (2.71).

Comparison group

The comparison group implemented *The Creative Curriculum*® without *Ladders to Literacy*. *The Creative Curriculum*® is a comprehensive curriculum for 3- to 5-year-old children. It addresses four areas of development: social/emotional, physical, cognitive, and language development. The curriculum required the physical space of the classroom to be structured into 10 interest areas: blocks, dramatic play, toys and games, art, library, discovery, sand and water, music and movement, cooking, and computers. Time was also allotted for outdoor activities. The 10 interest areas were designed to address curriculum content such as literacy, math, science, social studies, the arts, and technology, as well as process skills such as observing, exploring, and problem solving. *The Creative Curriculum*® included a developmental checklist that teachers were asked to use in ongoing assessments of child progress. Fidelity of implementation was assessed by conducting observations from December 2003 through April 2004 in the classrooms assigned to use *The Creative Curriculum*®. Researchers used a global fidelity measure to rate the overall fidelity with which the curriculum was implemented. On a four-point scale (0 = "not at all" to 3 = "high"), classrooms implementing *The Creative Curriculum*® were rated at the medium level (2.0).

Outcomes and measurement

The outcome domains of oral language, print knowledge, phonological processing, and math were assessed with standardized measures. Oral language was assessed with the PPVT-III and the Grammatic Understanding subtest from the TOLD-P:3. Print knowledge was assessed with the TERA-III and the WJ-III Letter-Word Identification and Spelling subtests. Phonological processing was assessed with the Elision subtest from the Pre-CTOPPP. Math was assessed with the WJ-III Applied Problems subtest, the composite score from the CMA-A, and the Building Blocks Shape Composition test. Pretesting was done in the fall of the preschool year, and posttesting was done in the spring of the preschool year. Trained research staff administered all assessments, which were conducted in English. For a more detailed description of these outcome measures, see Appendix B.

Support for implementation

All 14 teachers (both intervention and comparison) received at least one day of training on *The Creative Curriculum*® from a staff member at Teaching Strategies, Inc. Intervention group teachers received *Ladders to Literacy* training in September 2003, and ongoing training on a monthly basis throughout the 2003–04 school year. In addition to the September 2003 training, six intervention group teachers received training on *Ladders to Literacy* activities in September 2002.

Appendix B: Outcome measures for each domain

Oral language	
Mean length of utterance (MLU)	All utterances (at least 50 per child) were recorded for each child in the intervention and comparison groups. This required 15–30 minutes of recording per child. Utterances were transcribed verbatim by staff who were blind to the intervention status of the child. The median 50 utterances were selected from the resulting transcriptions and used to compute MLU. For a given child, MLU is calculated as the number of morphemes divided by the number of utterances, based on the entire sample of the child's speech (as cited in Russell, 2005).
Peabody Picture Vocabulary Test–III (PPVT-III)	A nationally-standardized, individually-administered assessment of children's receptive vocabulary for which children demonstrate understanding of a spoken word by pointing to a picture that best represents the meaning (as cited in PCER Consortium, 2008).
Test of Language Development–Primary III (TOLD-P:3) Grammatic Understanding subtest	A nationally-standardized, individually-administered assessment of children's ability to comprehend the meaning of sentences by selecting pictures that most accurately represent the sentence (as cited in PCER Consortium, 2008).
Type token ratio (TTR)	All utterances (at least 50 per child) were recorded for each child in the intervention and comparison groups. This required 15–30 minutes of recording per child. Utterances were transcribed verbatim by staff who were blind to the intervention status of the child. The median 50 utterances were selected from the resulting transcriptions and used to compute TTR. For a given child, TTR is calculated as the number of different words in the sample divided by the total number of words in the sample, based on the entire sample of the child's speech (as cited in Russell, 2005).
Print knowledge	
Test of Early Reading Ability–III (TERA-3) total score	A nationally-standardized, individually-administered assessment of children's developing reading skills with three subtests: alphabet, conventions, and meaning (as cited in PCER Consortium, 2008).[8]
Woodcock-Johnson III (WJ-III) Letter-Word Identification subtest	A nationally-standardized, individually-administered measure of identification of letters and reading of words (as cited in PCER Consortium, 2008).
WJ-III Spelling subtest	A nationally-standardized, individually-administered assessment that measures children's prewriting skills, such as drawing lines, tracing, and writing letters (as cited in PCER Consortium, 2008).
Phonological processing	
Preschool Comprehensive Test of Phonological and Print Processing (Pre-CTOPPP) Elision subtest	An individually-administered assessment of children's ability to identify and manipulate sounds in spoken words, using word prompts and picture plates for the first nine items and word prompts only for later items (as cited in PCER Consortium, 2008).
Math	
Building Blocks Shape Composition task	An individually-administered assessment of early math achievement, this measure was modified for PCER from the Early Maths Assessment, developed by Clements, Sarama, and Liu (2008). Children use blocks to fill in a puzzle and are assessed on whether they fill the puzzle without gaps or hangovers (as cited in PCER Consortium, 2008).
Child Math Assessment–Abbreviated (CMA-A) Composite score	An individually-administered assessment of early math achievement, this measure is the average of four subscales: (a) solving addition and subtraction problems using visible objects, (b) constructing a set of objects equal in number to a given set, (c) recognizing shapes, and (d) copying a pattern using objects that vary in color and identity from the model pattern. This assessment was adapted for PCER from a more comprehensive early math assessment by Klein and Starkey (2002), who also developed the pre-K math curriculum and participated in one of the research teams for PCER (as cited in PCER Consortium, 2008).
WJ-III Applied Problems subtest	A nationally-standardized, individually-administered assessment of children's ability to solve numerical and spatial problems, presented verbally with accompanying pictures of objects (as cited in PCER Consortium, 2008).

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

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			*p*-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
Russell, 2005[a]								
Mean length of utterance (MLU)	Preschool children	12 classrooms/ 34 children	3.36 (0.56)	3.45 (0.80)	–0.09	–0.13	–5	0.31
Type token ratio (TTR)	Preschool children	12 classrooms/ 34 children	0.52 (0.49)	0.50 (0.06)	0.02	0.05	+2	0.09
Domain average for oral language (Russell, 2005)						**–0.04**	**–1**	**Not statistically significant**
PCER Consortium, 2008[b]								
Peabody Picture Vocabulary Test III (PPVT-III)	Preschool children	14 classrooms/ 104 children	88.24 (18.03)	95.43 (14.88)	–7.19	–0.38	–15	> 0.05
Test of Language Development-–Primary III (TOLD-P:3) Grammatic Understanding subtest	Preschool children	14 classrooms/ 105 children	8.38 (2.87)	9.45 (2.61)	–1.07	–0.22	–9	> 0.05
Domain average for oral language (PCER Consortium, 2008)						**–0.30**	**–12**	**Not statistically significant**
Domain average for oral language across all studies						**–0.17**	**–7**	**na**

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

[a] For Russell (2005), corrections for clustering and multiple comparisons were needed but did not affect whether any of the contrasts were found to be statistically significant. The *p*-values presented here were reported in the original study. This study is characterized as having an indeterminate effect because the mean effect is neither statistically significant nor substantively important, accounting for multiple comparisons.

[b] For PCER Consortium (2008), corrections for multiple comparisons were needed but did not affect whether any of the contrasts were found to be statistically significant. The effect sizes, mean differences, and *p*-values presented here were reported in the original study (in Table 4.4, Table D-4a, and Table 4.4, respectively). Adjustment for the baseline pretest scores was not required for this domain. Thus, the intervention group mean equals the sum of the unadjusted comparison group mean and the covariate-adjusted mean difference reported in the study (in Table C-4a and Table D-4a, respectively). This study is characterized as having a potentially negative effect because the mean effect is substantively important and negative, accounting for multiple comparisons.

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

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			*p*-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
PCER Consortium, 2008[a]								
Test of Early Reading Ability–III (TERA-3) total score	Preschool children	14 classrooms/ 105 children	nr	nr	nr	–0.30	–12	> 0.05
Woodcock-Johnson III (WJ-III) Letter-Word Identification subtest	Preschool children	14 classrooms/ 105 children	nr	nr	nr	0.04	+2	> 0.05
WJ-III Spelling subtest	Preschool children	14 classrooms/ 105 children	nr	nr	nr	0.27	+11	> 0.05
Domain average for print knowledge (PCER Consortium, 2008)						**–0.00**	**0**	**Not statistically significant**
Domain average for print knowledge across all studies						**–0.00**	**0**	**na**

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

[a] For PCER Consortium (2008), corrections for multiple comparisons were needed but did not affect whether any of the contrasts were found to be statistically significant. The effect sizes and *p*-values presented here were reported in the original study (in Table A-10, based on an alternative estimation approach, ANCOVA, that included the baseline pretest). Mean scores and differences are not reported in this table because the study-reported group means and differences were not adjusted for the baseline pretest scores. This study is characterized as having an indeterminate effect because the mean effect is neither statistically significant nor substantively important, accounting for multiple comparisons.

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

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			*p*-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
PCER Consortium, 2008[a]								
Preschool Comprehensive Test of Phonological and Print Processing (Pre-CTOPPP) Elision subtest	Preschool children	14 classrooms/ 105 children	8.55 (4.60)	9.10 (5.09)	–0.55	–0.16	–6	> 0.05
Domain average for phonological processing (PCER Consortium, 2008)						**–0.16**	**–6**	**Not statistically significant**
Domain average for phonological processing across all studies						**–0.16**	**–6**	**na**

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

[a] For PCER Consortium (2008), the effect size, mean difference, and *p*-value presented here were reported in the original study (in Table 4.4, Table D-4a, and Table 4.4, respectively). Adjustment for the baseline pretest scores was not required for this domain. The WWC calculated the intervention group mean by adding the difference-in-differences adjusted estimate of the average impact of the program (i.e., difference in mean gains between the intervention and comparison groups) to the unadjusted comparison group posttest means reported in the study (in Table C-4a and Table D-4a, respectively). Please see the WWC Procedures and Standards Handbook, version 2.1, p. 96 for more information. This study is characterized as having an indeterminate effect because the single effect is neither statistically significant nor substantively important.

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

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			*p*-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
PCER Consortium, 2008[a]								
Building Blocks Shape Composition task	Preschool children	14 classrooms/ 104 children	nr	nr	nr	0.10	+4	> 0.05
Child Math Assessment–Abbreviated (CMA-A) Composite score	Preschool children	14 classrooms/ 105 children	nr	nr	nr	0.11	+4	> 0.05
Woodcock-Johnson III (WJ-III) Applied Problems subtest	Preschool children	14 classrooms/ 105 children	nr	nr	nr	0.03	+1	> 0.05
Domain average for math (PCER Consortium, 2008)						**0.08**	**+3**	**Not statistically significant**
Domain average for math across all studies						**0.08**	**+3**	**na**

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

[a] For PCER Consortium (2008), corrections for multiple comparisons were needed but did not affect whether any of the contrasts were found to be statistically significant. The effect sizes and *p*-values presented here were reported in the original study (in Table A-10, based on an alternative estimation approach, ANCOVA, that included the baseline pretest). Mean scores and differences are not reported in this table because the study-reported group means and differences were not adjusted for the baseline pretest scores. This study is characterized as having an indeterminate effect because the mean effect is neither statistically significant nor substantively important, accounting for multiple comparisons.

Endnotes

[1] The descriptive information for this program was obtained from a publicly available source: the program's website (http://products.brookespublishing.com/Ladders-to-Literacy-P557.aspx, downloaded November 2012). The WWC requests distributors review the program description sections for accuracy from their perspective. The program description was provided to the distributor in March 2012 and the WWC incorporated feedback from the developer. Further verification of the accuracy of the descriptive information for this program is beyond the scope of this review. The literature search reflects documents publicly available by December 2012.

[2] The previous report was released in August 2010. This report has been updated to include reviews of five studies that have been released since August 2010. None of the additional studies were within the scope of the review protocol for the Early Childhood Education topic area. A complete list and disposition of all studies reviewed are provided in the references. The report includes reviews of all previous studies that met WWC evidence standards with or without reservations and confirmed the study disposition of meets standards without reservations for Russell (2005) and the study disposition of meets standards with reservations for PCER Consortium (2008). The studies in this report were reviewed using the Evidence Standards from the WWC Procedures and Standards Handbook (version 2.1), along with those described in the Early Childhood Education topic area protocol (version 2.0). The evidence presented in this report is based on available research. Findings and conclusions may change as new research becomes available.

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

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

[5] These two studies examine the effects of *Ladders to Literacy* based on data from the PCER initiative. However, each study examines a unique sample of preschool children and examines different outcome measures. Russell (2005) examines the effects of that intervention for a sample of children attending preschool classrooms during the pilot year (first year) of the PCER initiative. PCER Consortium (2008) examines the effects of *Ladders to Literacy* for a sample of children attending preschool classrooms during the second year of the PCER initiative. Nine teachers participated in both the pilot year and the second year of the initiative, but five new teachers were included in the study only for the second year. The average age of the children in the sample examined in Russell (2005) and PCER Consortium (2008) is 4.7 years and 4.6 years, respectively.

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

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

[8] By name, this measure sounds as if it should be captured under the early reading and 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.

Recommended Citation

U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse. (2013, March). *Early Childhood Education intervention report: Ladders to Literacy*. Retrieved from http://whatworks.ed.gov.

WWC Rating Criteria

Criteria used to determine the rating of a study

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

Criteria used to determine the rating of effectiveness for an intervention

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

Criteria used to determine the extent of evidence for an intervention

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

Glossary of Terms

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

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

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

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

Domain A domain is a group of closely related outcomes.

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

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

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

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

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

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

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

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

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

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

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

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

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

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