What Works Clearinghouse

INSTITUTE OF EDUCATION SCIENCES

Early Childhood Education October 12, 2006

Dialogic Reading

Practice description

Dialogic Reading is an interactive shared picture book reading practice designed to enhance young children's language and literacy skills. During the shared reading practice, the adult and the child switch roles so that the child learns to become the storyteller with the assistance of the adult, who functions as an active listener and questioner. Two related practices are reviewed in the WWC intervention reports on *Interactive Shared Book Reading* and *Shared Book Reading*.

Research[1]

Four studies of *Dialogic Reading* met the What Works Clearinghouse (WWC) evidence standards and two studies met the WWC evidence standards with reservations.[2] Together, these six studies included more than 481 preschool children and examined intervention effects on children's oral language, print knowledge, phonological processing, and early reading/writing. Most of the children studied were from economically disadvantaged families. This report focuses on immediate posttest findings to determine the effectiveness of the intervention; however, follow-up findings provided by the study authors are included in the technical appendices.

Dialogic Reading was found to have positive effects on oral language, potentially positive effects on print knowledge, no discernible effects on phonological processing, and potentially positive effects on early reading/writing.

Effectiveness

	Oral language[3]	*Print knowledge*	*Phonological processing*	*Early reading/ writing*	*Cognition*	*Math*
Rating of effectiveness	Positive effects	Potentially positive effects	No discernible effects	Potentially positive effects	Not reported	Not reported
Improvement index[4]	*Average:* +17 percentile points *Range:* –12 to +48 percentile points	*Average:* +24 percentile points	*Average:* +5 percentile points *Range:* –7 to +40 percentile points	*Average:* +20 percentile points	Not reported	Not reported

1. The evidence presented in this report is based on available research. Findings and conclusions may change as new research becomes available.
2. To be eligible for the WWC's review, the Early Childhood Education (ECE) interventions had to be implemented in English in center-based settings with children ages 3 to 5 or in preschool.
3. In this domain, the Lonigan and Whitehurst (1998) study provided findings separately for high and low compliance centers. The rating of effectiveness is based on the effects from the high and low compliance centers combined, but findings are described separately for the high and low compliance centers in the findings section and appendix A4.
4. These numbers show the average and the range of improvement indices for all findings across the studies.

Absence of conflict of interest

The WWC ECE topic team works with two principal investigators: Dr. Ellen Eliason Kisker and Dr. Christopher Lonigan. The studies on *Dialogic Reading* reviewed by the ECE team included a number of studies on which Dr. Lonigan was either the primary or a secondary author and a number of studies on which Dr. Grover Whitehurst (Director, Institute for Education Sciences) was either a primary or a secondary author. Drs. Lonigan and Whitehurst's financial interests are not affected by the success or failure of *Dialogic Reading*, and they do not receive any royalties or other monetary return from the use of *Dialogic Reading*. In all instances where Drs. Lonigan and Whitehurst were study authors, they were not involved in the decision to include the study in the review, and they were not involved in the coding, reconciliation, or discussion of the included study. Dr. Kisker led all review activities related to those studies. The decision to review *Dialogic Reading* was made by Dr. Kisker, as co-principal investigator, in collaboration with the rest of the ECE team following prioritization of interventions based on the results from the literature review. This report on *Dialogic Reading* was reviewed by a group of independent reviewers, including members of the WWC Technical Review Team and external peer reviewers.

Additional practice information

Developer and contact

Dialogic Reading is a practice that does not have a single developer responsible for providing information or materials. A non-comprehensive list of sources of the *Dialogic Reading* materials, not reviewed or endorsed by the WWC, include:

- Pearson Early Learning: www.pearsonearlylearning.com/products/curriculum/rttt/index.html;
- The Committee for Children: www.cfchildren.org/wwf/dialogic;
- Rotary Club of Bainbridge Island in Washington State: www.bainbridgeislandrotary.org/default.aspx?c=10052;
- Reading Rockets: www.pbs.org/launchingreaders/rootsofreading/meettheexperts_2.html;
- The American Library Association: www.ala.org/ala/alsc/alscresources/borntoread/bornread.htm.

Scope of use

Dialogic Reading was created in the 1980s, and the first published study appeared in 1988 (Whitehurst, Falco, Lonigan, Fischel, DeBaryshe, Valdez-Menchaca, & Caulfield, 1988).[5] Information is not available on the number or demographics of children or centers using this intervention.

Teaching

In center-based settings, *Dialogic Reading* can be used by teachers with children individually or in small groups. Teachers can be trained on the principles of *Dialogic Reading* through videotape followed by role-playing and group discussion.

While reading books with the child, the adult uses five types of prompts (CROWD):

- **C**ompletion: child fills in blank at the end of a sentence.
- **R**ecall: adult asks questions about a book the child has read.
- **O**pen-ended: adult encourages child to tell what is happening in a picture.
- **W**h-: adult asks "wh-" questions about the pictures in books.
- **D**istancing: adult relates pictures and words in the book to children's own experiences outside of the book.

These prompts are used by the adult in a reading technique called PEER:

5. Whitehurst, G. J., Falco, F. L., Lonigan, C. J., Fischel, J. E., DeBaryshe, B. D., Valdez-Menchaca, M. C., & Caulfield, M. (1988). Accelerating language development through picture book reading. *Developmental Psychology*, *24*(4), 552–559. This study was not reviewed because it fell outside the scope of the current ECE review (that is, the study was not center-based and children were younger than 3 years old).

Additional practice information *(continued)*

- **P**: adult prompts the child to say something about the book.
- **E**: adult evaluates the response.
- **E**: adult expands the child's response.
- **R**: adult repeats the prompt.

As the child becomes increasingly familiar with a book, the adult reads less, listens more, and gradually uses more higher level prompts to encourage the child to go beyond naming objects in the pictures to thinking more about what is happening in the pictures and how this relates to the child's own experiences.

Cost

Published *Dialogic Reading* procedures are freely available to the public. Information is not available about the costs of teacher training and implementation of *Dialogic Reading*.

Research

Eight studies reviewed by the WWC investigated the effects of *Dialogic Reading* in center-based settings. Four studies (Lonigan, Anthony, Bloomfield, Dyer, & Samwel, 1999; Lonigan & Whitehurst, 1998; Wasik & Bond, 2001; Whitehurst, Epstein, Angell, Payne, Crone, & Fischel, 1994[6]) were randomized controlled trials that met WWC evidence standards. Two studies (Crain-Thoreson & Dale, 1999; Whitehurst, Arnold, Epstein, Angell, Smith, & Fischel, 1994) were randomized controlled trials with differential attrition that lowered the evidence ratings of the studies to met WWC evidence standards with reservations. The remaining two studies did not meet WWC evidence screens.

Met evidence standards

Lonigan et al. (1999) included 95 two- to five-year-old predominantly low-income children from five child care centers in an urban area in Florida. Lonigan et al. compared two interventions—*Dialogic Reading* and typical shared book reading—to a no-treatment comparison group. This report focuses on the comparison of oral language and phonological processing outcomes between the *Dialogic Reading* group[7] and the no-treatment comparison group with a total of 66 children.

Lonigan and Whitehurst (1998) included 91 low-income three- to four-year-old children from four child care centers in Nashville, Tennessee. Lonigan and Whitehurst compared three intervention groups—*Dialogic Reading* at school, *Dialogic Reading* at home, and *Dialogic Reading* both at school and at home—to a no-treatment comparison group. This report focuses on the comparison of oral language outcomes between the combined school and school plus home group and the no-treatment comparison groups[8] with a total of 75 children.

Wasik and Bond (2001) included 121 low-income three- to four-year-old children from a Title I early learning center in Baltimore, Maryland. Wasik and Bond compared oral language outcomes for children participating in *Dialogic Reading* plus reinforcement activities with outcomes for children in a comparison condition who were read the same books by teachers with no training in *Dialogic Reading*.

Whitehurst, Epstein, et al. (1994) included 167 at-risk low-income four-year-old children from four Head Start centers in Suffolk County, New York. Whitehurst, Epstein, et al. compared oral language, phonological processing, print knowledge, and early reading/writing outcomes for children participating in *Dialogic Reading* combined with an adapted Sound Foundations cur-

6. Zevenbergen, Whitehurst, & Zevenbergen (2003) reports additional results from the study first reported in Whitehurst, Epstein, et al. (1994) and was reviewed along with that study.
7. The comparison between the typical shared book reading group and the comparison group is included in the WWC *Shared Book Reading* intervention report.
8. The *Dialogic Reading* at home group is not included in the review because it is not center-based. The *Dialogic Reading* at school and the *Dialogic Reading* at school and at home groups were combined for this review to reflect analyses conducted by the study authors. However, the data separated for these two groups are included in Appendix A4.1. The study authors divided centers into high- and low-compliance centers based on the frequency level of *Dialogic Reading* sessions. The WWC report includes findings for the high and low compliance centers combined in the overall rating of effectiveness, and describes findings separated by high and low compliance in the findings section and appendix A4.

Research *(continued)*

riculum to outcomes for children participating in a no-treatment comparison group.

Met evidence standards with reservations

Crain-Thoreson and Dale (1999) included 32 three- to five-year-old children with mild to moderate language delays from five classrooms in three school districts in the Pacific Northwest. Crain-Thoreson and Dale compared two intervention groups—a staff-implemented *Dialogic Reading* group (staff/practice) and a parent-implemented *Dialogic Reading* group (parent/practice)—to a comparison group that did not receive one-on-one *Dialogic Reading*. This report focuses on the comparison of oral language outcomes between the staff/practice group[9] and the comparison group with a total of 22 children.

Whitehurst, Arnold, et al. (1994) included 67 low-income three-year-old children from five day care centers in Suffolk County, New York. Whitehurst, Arnold, et al. compared two intervention groups—*Dialogic Reading* at school and *Dialogic Reading* both at school and at home—to a comparison group that participated in small-group play activities. This report focuses on the comparison of oral language outcomes between the combined school and school plus home group and the comparison group.[10]

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/writing, cognition, and math.[11]

Oral language. Six studies examined outcomes in the domain of oral language: three studies showed statistically significant and positive effects and three studies showed indeterminate effects.

Lonigan et al. (1999) found a statistically significant difference favoring children in the *Dialogic Reading* intervention group on one of the four outcome measures (verbal expression subscale of the Illinois Test of Psycholinguistic Ability; ITPA-VE), and this effect was confirmed to be statistically significant by the WWC. The authors found no statistically significant differences between the intervention and comparison groups on the other three measures. In this study, *Dialogic Reading* had a statistically significant and positive effect on oral language, according to WWC criteria.

Lonigan and Whitehurst (1998) analyzed group differences for the combined intervention groups (*Dialogic Reading* at school, both at school and at home, and at home) and the comparison group. Because WWC ECE does not review interventions implemented in the home, the WWC calculated group differences on the three outcome measures for the combined *Dialogic Reading* at school and both at school and at home intervention group versus the comparison group and did not find statistically significant differences on any measure in analyses using data combined for centers with high and low implementation. In this study, *Dialogic Reading* had an indeterminate effect on oral language, according to WWC criteria.

Lonigan and Whitehurst (1998) also analyzed group differences for the combined intervention groups within high and low compliance centers. The WWC calculated group differences on the three outcome measures for the combined *Dialogic Reading*

9. The parent/practice group was not included in the review because it was not center-based.
10. The *Dialogic Reading* at school and the *Dialogic Reading* at school and at home groups were combined for this review to reflect analyses conducted by the study authors. However, the data separated for these two groups are included in Appendix A4.
11. The level of statistical significance was reported by the study authors or, where necessary, calculated by the WWC to correct for clustering within classrooms or schools and for multiple comparisons. For an explanation about the clustering correction, see the WWC Tutorial on Mismatch. See Technical Details of WWC-Conducted Computations for the formulas the WWC used to calculate the statistical significance. In the case of *Dialogic Reading*, corrections for clustering and multiple comparisons were needed.

at school and both at school and at home intervention group versus the comparison group separately for high and low compliance centers. For the high compliance centers, the WWC did not find statistically significant differences on any measure; however, for one measure the effect was large enough to be called substantively important. For high compliance centers, *Dialogic Reading* had a substantively important and positive effect on oral language, according to WWC criteria. For low compliance centers, the WWC did not find statistically significant differences on any measure. For low compliance centers, *Dialogic Reading* had an indeterminate effect on oral language, according to WWC criteria. These analyses suggest that the level of implementation of *Dialogic Reading* has an impact on child outcomes in the oral language domain.

Wasik and Bond (2001) found statistically significant differences favoring the *Dialogic Reading* children on two measures of oral language, and the WWC confirmed this statistical significance.[12] In this study, *Dialogic Reading* had a statistically significant and positive effect on oral language, according to WWC criteria.

Whitehurst, Epstein, et al. (1994) found no statistically significant difference between the intervention group and the comparison group on oral language as measured by the Language factor.[13] Zevenbergen, Whitehurst, and Zevenbergen (2003) reported findings on four additional oral language measures from the same study, none of which were statistically significant as calculated by the WWC. The average effect of the five measures across the two study reports was not large enough to be considered substantively important. In this study, the effect of *Dialogic Reading* on oral language was indeterminate, according to WWC criteria.

Crain-Thoreson and Dale (1999) analyzed findings for six measures in this outcome domain. The findings favored the intervention group for five of the measures and favored the comparison group for the sixth measure. None of these effects, however, were statistically significant, and the average effect was neither statistically significant nor large enough to be considered substantively important. In this study, the effect of *Dialogic Reading* on oral language was indeterminate, according to WWC criteria.

Whitehurst, Arnold, et al. (1994) found statistically significant differences favoring children in the *Dialogic Reading* intervention group on two of the four measures in this domain (EOWPVT-R and Our Word), but only the statistical significance for EOWPVT-R was confirmed by the WWC. The authors found no statistically significant differences on the other two measures.[14] In this study, *Dialogic Reading* was considered to have a statistically significant and positive effect on oral language, according to WWC criteria.

Print knowledge. Whitehurst, Epstein, et al. (1994) reported, and the WWC confirmed, a statistically significant difference favoring the *Dialogic Reading* group on the Print Concepts factor.[13] So, in this study, *Dialogic Reading* had a statistically significant and positive effect on print knowledge, according to WWC criteria.

Phonological processing. Lonigan et al. (1999) found no statistically significant effects for any of the four outcome measures and the average effect across the four measures was not large enough to be considered substantively important. In this study, the effect of *Dialogic Reading* on phonological processing was indeterminate, according to WWC criteria. Whitehurst, Epstein, et al. (1994) reported neither statistically significant nor substantively important effects on the Linguistic Awareness factor.[13] In this study, the effect of *Dialogic Reading* on phonological processing was indeterminate, according to WWC criteria.

Early reading/writing. Whitehurst, Epstein, et al. (1994) reported a statistically significant difference favoring the *Dialogic Reading* group on the Writing Factor.[13] This effect was confirmed

12. The authors also reported findings on the Peabody Picture Vocabulary Test-III (PPVT-III), but there was not enough information to compute an effect size, so this measure was not included in the review.
13. The study authors conducted a principal components analysis on the 21 measures to reduce data. The WWC presents results only for the factor score because effect sizes could not be computed for the individual measures.
14. The authors also reported results from the six-month follow-up tests. Since the primary focus of this review is on the immediate posttest results, the follow-up results are not discussed here but are included in Appendix A4.1.

Effectiveness *(continued)*

to be statistically significant by the WWC. Thus, in this study, *Dialogic Reading* had a statistically significant and positive effect on early reading/writing, according to WWC criteria.

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,[11] 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 *Dialogic Reading* to have positive effects for oral language, potentially positive effects for print knowledge, no discernible effects for phonological processing, and potentially positive effects for early reading/writing

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 entirely based on the size of the effect, regardless of the statistical significance of the effect, the study design, or the analysis. The improvement index can take on values between -50 and +50, with positive numbers denoting favorable results. The average improvement index for oral language is +17 percentile points across the six studies, with a range of -12 to +48 percentile points across findings. The improvement index for print knowledge is +24 percentile points for the one outcome in the Whitehurst, Epstein, et al. (1994) study. The average improvement index for phonological processing is +5 percentile points across the two studies, with a range of -7 to +40 percentile points across findings. The improvement index for early reading/writing is +20 percentile points for the one outcome in the Whitehurst, Epstein, et al. (1994) study.

Summary

The WWC reviewed eight studies on *Dialogic Reading*. Four of the studies met WWC standards; two studies met WWC standards with reservations, and the remaining studies did not meet WWC evidence screens. Based on these six studies, the WWC found positive effects for oral language, potentially positive effects for print knowledge, no discernible effects for phonological processing, and potentially positive effects for early reading/writing. The evidence presented in this report may change as new research emerges.

References

Met WWC evidence standards

Lonigan, C. J., Anthony, J. L., Bloomfield, B. G., Dyer, S. M., & Samwel, C. S. (1999). Effects of two shared-reading interventions on emergent literacy skills of at-risk preschoolers. *Journal of Early Intervention, 22*(4), 306–322.

Lonigan, C. J., & Whitehurst, G. J. (1998). Relative efficacy of parent and teacher involvement in a shared-reading intervention for preschool children from low-income backgrounds. *Early Childhood Research Quarterly, 13*(2), 263–290.

Wasik, B. A., & Bond, M. A. (2001). Beyond the pages of a book: Interactive book reading and language development in preschool classrooms. *Journal of Educational Psychology, 93*(2), 243–250.

Whitehurst, G. J., Epstein, J. N., Angell, A. L., Payne, A. C., Crone, D. A., & Fischel, J. E. (1994). Outcomes of an emergent literacy intervention in Head Start. *Journal of Educational Psychology, 86*(4), 542–555.

Additional source:

Zevenbergen, A. A., Whitehurst, G. J., & Zevenbergen, J. A. (2003). Effects of a shared-reading intervention on the inclusion of evaluative devices in narratives of children from low-income families. *Journal of Applied Developmental Psychology, 24*, 1–15.

References *(continued)*

Met WWC evidence standards with reservations

Crain-Thoreson, C., & Dale, P. S. (1999). Enhancing linguistic performance: Parents and teachers as book reading partners for children with language delays. *Topics in Early Childhood Special Education, 19*(1), 28–39.

Whitehurst, G. J., Arnold, D. S., Epstein, J. N., Angell, A. L., Smith, M., & Fischel, J. E. (1994). A picture book reading intervention in day care and home for children from low-income families. *Developmental Psychology, 30*(5), 679–689.

Did not meet WWC evidence screens

Hargrave, A. C., & Sénéchal, M. (2000). A book reading intervention with preschool children who have limited vocabularies: The benefits of regular reading and dialogic reading. *Early Childhood Research Quarterly, 15*(1), 75–90.[15]

Whitehurst, G. J., Zevenbergen, A. A., Crone, D. A., Schultz, M. D., Velting, O. N., & Fischel, J. E. (1999). Outcomes of an emergent literacy intervention from Head Start through second grade. *Journal of Educational Psychology, 91*(2), 267–272.[16]

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

15. Confound: there was only one cluster (childcare center) in each study condition; therefore, the effects of the intervention could not be separated from the effects of the cluster.
16. Complete data were not reported: the WWC could not compute effect sizes based on the data reported.

Appendix

Appendix A1.1 Study characteristics: Lonigan, Anthony, Bloomfield, Dyer, & Samwel, 1999 (randomized controlled trial)

Characteristic	Description
Study citation	Lonigan, C. J., Anthony, J. L., Bloomfield, B. G., Dyer, S. M., & Samwel, C. S. (1999). Effects of two shared-reading interventions on emergent literacy skills of at-risk preschoolers. *Journal of Early Intervention, 22*(4), 306–322.
Participants	The study began with 110 children; 15 children left the child care centers, leaving a sample of 95 children. Most of the children were from low-income families. The mean age of the child participants was 45.1 months (range 25 to 64 months). Forty-six percent were female and 77% were African-American. Results for the 66 children who had been randomly assigned within center to the *Dialogic Reading* and no-treatment comparison conditions are included in this report.
Setting	The study took place in five child care centers in an urban area in Florida. Four centers served primarily children of families eligible for subsidized child care. The fifth center was affiliated with a church and approximately 25% of families served by the church center received a state child care subsidy.
Intervention	The study included two intervention groups: *Dialogic Reading* and typical shared book reading. The *Dialogic Reading* intervention is included in this review; results involving typical shared book reading are included in the WWC *Shared Book Reading* report. In the *Dialogic Reading* condition, trained undergraduate volunteers engaged in *Dialogic Reading* intervention sessions for 10–15 minutes each day across a six-week period. Children were read to in small groups of three to five children in a location outside the classroom.
Comparison	Children in the no-treatment comparison group engaged in their standard preschool curriculum.
Primary outcomes and measurement	The primary outcome domains were children's oral language and phonological processing. The study used the following standardized measures: the Peabody Picture Vocabulary Test-Revised (PPVT-R), the Expressive One-Word Picture Vocabulary Test-Revised (EOWPVT-R), the Verbal Expression subscale of the Illinois Test of Psycholinguistic Abilities (ITPA-VE), and the Listening Comprehension subtest of the Woodcock-Johnson Psychoeducational Battery (WJ-LC). The study also utilized four measures of phonological processing: rhyme oddity detection, alliteration oddity detection, sound blending, and sound elision. (See Appendices A2.1 and A2.3 for more detailed descriptions of these outcome measures.)
Teacher training	Undergraduate volunteer readers were trained in *Dialogic Reading* using a videotape training method, which covered the two phases of *Dialogic Reading*. During the training, the trainees were presented with *Dialogic Reading* guidelines and watched vignettes of adult-child shared book reading that followed or did not follow the guidelines. Trainees analyzed the vignettes and had one-on-one role plays with the trainer. The phase one training session lasted 30 minutes, the phase two session 20 minutes.

Appendix A1.2 Study characteristics: Lonigan & Whitehurst, 1998 (randomized controlled trial)

Characteristic	Description
Study citation	Lonigan, C. J., & Whitehurst, G. J. (1998). Relative efficacy of parent and teacher involvement in a shared-reading intervention for preschool children from low-income backgrounds. *Early Childhood Research Quarterly, 13*(2), 263–290.
Participants	The study began with 114 children from low-income households; 23 of these children left the child care center they were attending prior to the posttest, leaving 91 children in the sample. These 91 children were between 33 and 60 months of age at the time of pretest. Fifty-four percent were female, 91% were African-American, and all children were from English-speaking homes. The children were randomly assigned within classroom to the intervention and comparison conditions.[1] Results for the 75 children who had been randomly assigned to the *Dialogic Reading* at school, *Dialogic Reading* both at school and at home, and comparison groups are included in this report.
Setting	The study took place in four child care centers in Nashville, Tennessee, that served primarily children of families eligible for subsidized child care.
Intervention	The study included three intervention groups: *Dialogic Reading* at school, *Dialogic Reading* at home, and *Dialogic Reading* both at school and at home. The *Dialogic Reading* at home group is not included in this review because it is not center-based. The *Dialogic Reading* at school and the *Dialogic Reading* both at school and at home groups were combined for this review to reflect analyses conducted by the study authors. *Dialogic Reading* was implemented over a six-week period. Teachers or aides conducted *Dialogic Reading* sessions with children in small groups of fewer than six children. Sessions were planned to take place every day for about 10 minutes. The study authors divided centers into low and high compliance centers based on the frequency level (i.e., high and low) of *Dialogic Reading* sessions. The WWC uses the findings for the high and low compliance centers combined to determine the overall rating of effectiveness; however, the WWC reports findings for the low and high compliance centers separately in Appendix A4.
Comparison	Children in the no-treatment comparison group did not participate in *Dialogic Reading* at home or at school.
Primary outcomes and measurement	The primary outcome domain was children's oral language use. Standardized tests included the PPVT-R, the EOWPVT-R, and the ITPA-VE. Lonigan and Whitehurst also included measures of verbal production (MLU, speech production, diversity, and semantic diversity) that are not included in this review because of attrition. (See Appendix A2.1 for more detailed descriptions of these outcome measures.)
Teacher training	Teachers were trained in *Dialogic Reading* using a videotape training method that covered the two phases of *Dialogic Reading*. During the training, the trainees were presented with *Dialogic Reading* guidelines and watched vignettes of adult-child shared book reading on tape that followed or did not follow the guidelines. Trainees analyzed the vignettes and had one-on-one role plays with the trainer. The phase one training session lasted 30 minutes, the phase two session 20 minutes.

1. Although the authors did some reassignment of children where necessary to obtain equality in pretest scores, the WWC did not downgrade the study, because reassignment of children was minimal and resulted in a more conservative test of the intervention effects.

Appendix A1.3 Study characteristics: Wasik & Bond, 2001 (randomized controlled trial)

Characteristic	Description
Study citation	Wasik, B. A., & Bond, M. A. (2001). Beyond the pages of a book: Interactive book reading and language development in preschool classrooms. *Journal of Educational Psychology, 93*(2), 243–250.
Participants	The study began with 127 children from low-income households ranging in age from 3.9 years to 4.6 years (mean age = 4.3 years). After assignment to groups, six children transferred to another school, leaving 121 children in the sample. The center that the children attended served primarily three- to five-year-old children who were eligible for free or reduced-price lunch. Among the children attending the center, 94% were African-American. Four teachers were randomly assigned to the intervention and comparison conditions.
Setting	The study took place in a Title I early learning center in Baltimore, Maryland.
Intervention	The study examined the effect of interactive book reading plus extension activities reinforcing the use of target vocabulary in the book on children's language development. The interactive reading sessions were conducted by trained teachers with the whole class of children, and the extension activities supporting the use of target vocabulary were conducted in various contexts. The intervention took place four days a week and lasted for 15 weeks in the children's classrooms. Two books a week were read; one book was read twice and the other book was read once.
Comparison	The comparison group teachers read the same books the same number of times as the intervention group. However, the comparison group teachers were not trained to use the interactive book reading strategies.
Primary outcomes and measurement	The primary outcome domain was children's oral language use. This domain was assessed with one standardized measure and two researcher-developed measures. The standardized measure was the PPVT-III. The researcher-developed measures were: a receptive language measure, which was developed using a subset of vocabulary words presented in the interactive book reading intervention, and an expressive language measure, which contained pictures representing the words presented during the interactive book reading intervention. The PPVT-III is not considered in this review because an effect size could not be calculated from the information provided. (See Appendix A2.1 for more detailed descriptions of these outcome measures.)
Teacher training	Teachers were trained in interactive book reading techniques and book reading extension activities by an experienced teacher trainer. During the training, teachers were instructed in how to use interactive book reading strategies (for example, defining vocabulary words, asking open-ended questions, and giving children an opportunity to talk and be heard). The teacher trainer modeled these strategies in the intervention classrooms and assisted extension activities in the first four weeks of the intervention. In addition, each intervention teacher was given books and materials focused on early childhood topics and themes that would be used in the book reading and extension activities.

Appendix A1.4 Study characteristics: Whitehurst, Epstein, Angell, Payne, Crone, & Fischel, 1994 and Zevenbergen, Whitehurst, & Zevenbergen, 2003 (randomized controlled trials)

Characteristic	Description
Study citations	Whitehurst, G. J., Epstein, J. N., Angell, A. L., Payne, A. C., Crone, D. A., & Fischel, J. E. (1994). Outcomes of an emergent literacy intervention in Head Start. *Journal of Educational Psychology, 86*(4), 542–555. ***Additional source:*** Zevenbergen, A. A., Whitehurst, G.J., & Zevenbergen, J. A. (2003). Effects of a shared-reading intervention on the inclusion of evaluative devices in narratives of children from low-income families. *Journal of Applied Developmental Psychology, 24*, 1–15.
Participants	The study began with 207 four-year-old at-risk low socioeconomic status children. Forty children did not complete the study, leaving 167 children in the final sample. The final sample of children was 46% Caucasian, 45% African-American, 8% Latin American, and 1% Asian, and 44% of the sample was female. Fifteen classrooms were randomly assigned to the intervention and comparison conditions.[1]
Setting	The study took place in 15 classrooms from four Head Start centers in Suffolk County, New York.
Intervention[2]	The study included two intervention elements: emergent literacy (*Dialogic Reading* plus an adaptation of Sound Foundations) at school and one-on-one *Dialogic Reading* at home. *Dialogic Reading* occurred over a 30-week period and consisted of reading to children in small groups three to five times a week in the classroom and one-on-one reading at home with the same book. A different book was used each week, and the researchers added hints to each book (for example, wh- and recall prompts). Sound Foundations occurred one to three times a week over a 16-week period in the classroom. Children were introduced to seven consonant sounds at the beginning and ending of words, two vowel sounds at the beginning of words, and manuscript letters that correspond to curriculum sounds.
Comparison	Children in the no-treatment comparison group participated in their regular "business as usual" Head Start program.
Primary outcomes and measurement	Whitehurst, Epstein, et al. (1994) examined outcomes in the oral language, phonological processing, print knowledge, and early reading/writing domains. Children's oral language was measured by three standardized measures: PPVT-R; EOWPVT-R; and ITPA-VE. Children's literacy was measured by 18 sub-scales from the Developing Skills Checklist. Because of the large number of outcome measures (21), the study authors conducted a principal components analysis to reduce the data to four factors for the analyses: language (oral language domain), print concepts (print knowledge domain), linguistic awareness (phonological processing domain), and writing (early reading/writing domain). (See Appendices A2.1–A2.4 for more detailed descriptions of these outcome measures.) Zevenbergen et al. (2003) tested additional outcomes from the same study. They assessed children's narrative understanding by asking children to retell a story about a bus immediately after hearing the story. Their narrative was transcribed and coded for general content (information) and children's use of narrative devices (references to character states, dialogue, and causal states, all in the oral language domain). (See Appendix A2.1 for more detailed descriptions of these outcome measures.)
Teacher training	Parents and teachers were trained on *Dialogic Reading* by the authors using a 20-minute video combined with role-playing and discussion after viewing the video. Training occurred once at the beginning of the school year. Teachers and aides in the intervention classrooms were asked to keep a daily log of the reading activities. To observe compliance and provide feedback, each classroom was visited at least once every two weeks by one of the study authors. Specific training for Sound Foundations is not reported.

1. Zevenbergen et al. (2003) reported that 16 classrooms were assigned randomly.
2. Note that the intervention in this study was a combination of both *Dialogic Reading* and *Sound Foundations*. This study is also included in the WWC *Sound Foundations* intervention report.

Appendix A1.5 Study characteristics: Crain-Thoreson & Dale, 1999 (randomized controlled trial with attrition problems)

Characteristic	Description
Study citation	Crain-Thoreson, C., & Dale, P. S. (1999). Enhancing linguistic performance: Parents and teachers as book reading partners for children with language delays. *Topics in Early Childhood Special Education, 19*(1), 28–39.
Participants	The study began with 37 children with mild to moderate language delays. All children were eligible for early childhood special education services. Five children did not complete the study, leaving 32 children remaining in the sample.[1] The mean age of the remaining children was 51.6 months (range 39 to 66 months) and 31.3% of the remaining sample was female. Results for the 22 children who were randomly assigned to the staff/practice and comparison conditions are included in this report.
Setting	The study took place in five classrooms in three school districts in the Pacific Northwest.
Intervention	The study included two intervention groups: a staff/practice group and a parent group. The staff/practice group is included in this review; the parent group was not included in the review because it was not center-based. *Dialogic Reading* occurred over an eight-week period during which staff engaged in book reading with individual children at least four times a week.
Comparison	Staff in the comparison group were trained on *Dialogic Reading*, but children did not participate individually in *Dialogic Reading*.
Primary outcomes and measurement	The primary outcome domain was children's oral language use as measured by four non-standardized measures: mean length of utterances, number of utterances, number of different words used, and ratio of child participation. Children's vocabulary knowledge was measured by two standardized tests: PPVT-R and the EOWPVT-R. (See Appendix A2.1 for more detailed descriptions of these outcome measures.)
Teacher training	Staff were trained on the *Dialogic Reading* program in two 1.5 hour instructional sessions held four weeks apart. Videotape training, live demonstration, and role-play were used to train teachers to use *Dialogic Reading*. In addition to learning *Dialogic Reading* principles, staff were instructed to pause and give children time to respond. They kept logs of their shared reading activities.

1. The study was downgraded by the WWC because of differential attrition between the intervention and comparison groups. The attrition was 0% for the staff/practice intervention group and 18% for the comparison group. The difference in attrition between groups was 18%.

Appendix A1.6 Study characteristics: Whitehurst, Arnold, Epstein, Angell, Smith, & Fischel, 1994 (randomized controlled trial with attrition problems)

Characteristic	Description
Study citation	Whitehurst, G. J., Arnold, D. S., Epstein, J. N., Angell, A. L., Smith, M., & Fischel, J. E. (1994). A picture book reading intervention in day care and home for children from low-income families. *Developmental Psychology, 30*(5), 679–689.
Participants	The study began with 73 three-year-old children from low-income families; at immediate posttest 67 children remained in the sample. At entry into the study, the mean age of the children was 3.5 years. Forty-five percent were female, 55% were African-American, and 23% were Hispanic. The children were randomly assigned within classroom to intervention and comparison conditions.[1]
Setting	The study took place in five day-care centers in Suffolk County, New York, which served mainly children of families qualified for public subsidy of day-care costs under Title XX of the Federal Social Security Act.
Intervention	The study included two intervention conditions: *Dialogic Reading* at school and *Dialogic Reading* both at school and at home. The *Dialogic Reading* at school and the *Dialogic Reading* both at school and at home groups were combined for this review to reflect analyses conducted by the study authors. In the *Dialogic Reading* at school condition, the teacher or aide conducted the sessions in the classroom in small groups of no more than five children daily for about 10 minutes over a six-week period. In the *Dialogic Reading* both at school and at home condition, children received similar small-group *Dialogic Reading* sessions at school and a one-on-one daily *Dialogic Reading* session at home with their parents.
Comparison	The comparison condition children participated in play activities in small groups of no more than five children daily for about 10 minutes. The play activities centered on construction toys that were not available in the classrooms before the study.
Primary outcomes and measurement	The primary outcome domain was children's oral language use. The study used the following standardized measures: the PPVT-R, the EOWPVT-R, and the ITPA-VE. The authors also utilized a researcher-developed measure called "Our Word." (See Appendix A2.1 for more detailed descriptions of these outcome measures.)
Teacher training	Teachers were trained in *Dialogic Reading* using a videotape training method, which presented the two phases of *Dialogic Reading*. During the training, the trainees were presented with a set of guidelines and taped vignettes of adult-child book reading that exemplified or did not follow the guidelines. Trainees critiqued the vignettes and had one-on-one role plays with the trainer. The two phases were presented three weeks apart, with the phase one training session lasting 30 minutes and the phase two session 20 minutes. Parents were trained to use *Dialogic Reading* at home using the same videotape and similar training procedures as the teachers at their child's day-care centers.

1. The authors did some reassignment of children where necessary to obtain equality in pretest scores. On its own, this reassignment of children did not lead the WWC to downgrade the study because reassignment of children was minimal and resulted in a more conservative test of the intervention effects. However, the study was downgraded because of differential attrition between the intervention and comparison groups. The attrition was 17% for the school plus home group and 8% for the comparison group. The difference in attrition between groups was 9%.

Appendix A2.1 Outcome measures in the oral language domain

Outcome measure	Description
Peabody Picture Vocabulary Test-Revised (PPVT-R)	A standardized measure of children's receptive vocabulary that requires them to identify pictures that correspond to spoken words (as cited in Crain-Thoreson & Dale, 1999; Lonigan & Whitehurst, 1998; Lonigan et al., 1999; Whitehurst, Arnold, et al., 1994).
Expressive One-Word Picture Vocabulary Test-Revised (EOWPVT-R)	A standardized measure of children's expressive vocabulary that requires them to name pictures of common objects, actions, and concepts (as cited in Crain-Thoreson & Dale, 1999; Lonigan & Whitehurst, 1998; Lonigan et al., 1999; Whitehurst, Arnold, et al., 1994).
Our Word	A researcher-developed measure designed to resemble the EOWPVT-R and measure children's knowledge of novel vocabulary introduced in books in the study (as cited in Whitehurst, Arnold, et al., 1994).
Illinois Test of Psycholinguistic Abilities-Verbal Expression subscale (ITPA-VE)	A standardized measure of children's verbal fluency that requires them to describe four common objects (as cited in Lonigan & Whitehurst, 1998; Lonigan et al., 1999; Whitehurst, Arnold, et al., 1994).
Number of child utterances	The number of child utterances during videotaped book sharing measured using the Computerized Language Analysis Programs (CLAN) (as cited in Crain-Thoreson & Dale, 1999).
Ratio of child participation	The number of child utterances divided by the sum of child and adult utterances during videotaped book sharing measured using the Computerized Language Analysis Programs (CLAN) (as cited in Crain-Thoreson & Dale, 1999).
Mean length of utterances	The mean length of utterances during videotaped book sharing measured using the Computerized Language Analysis Programs (CLAN) (as cited in Crain-Thoreson & Dale, 1999).
Lexical diversity	The number of different words spoken by the child during videotaped book sharing measured using the Computerized Language Analysis Programs (CLAN) (as cited in Crain-Thoreson & Dale, 1999).
Woodcock-Johnson Psychoeducational Battery-Listening Comprehension subtest (WJ-LC)	A standardized measure of children's listening comprehension that requires children to finish incomplete sentences by providing the missing word (as cited in Lonigan et al., 1999).
Language factor	A factor derived from a number of outcome measures (subscales from the Developing Skills Checklist, ITPA-VE, PPVT-R, and EOWPVT-R) using a principal components analysis. Factor loadings for the language factor were high for EOWPVT-R, PPVT-R, ITPA-VE, Tell a Story in Sequence, and Identify Function of Words-Numbers (as cited in Whitehurst, Epstein, et al., 1994).
Information/general content score	Children heard an adapted version of the Bus Story (Renfrew, 1969 as cited in Zevenbergen et al., 2003) and then retold the story. Researchers coded transcripts of the child's narrative to rate the general content of the child's story.
Reference to character states	Children heard an adapted version of the Bus Story (Renfrew, 1969 as cited in Zevenbergen et al., 2003) and then retold the story. Researchers coded transcripts of the child's narrative to determine the number of times the child referred to the internal states of the characters in the story.
Dialogue	Children heard an adapted version of the Bus Story (Renfrew, 1969 as cited in Zevenbergen et al., 2003) and then retold the story. Researchers coded transcripts of the children's narrative to determine their usage of dialogue.
Causal statements	Children heard an adapted version of the Bus Story (Renfrew, 1969 as cited in Zevenbergen et al., 2003) and then retold the story. Researchers coded transcripts of the children's narrative to determine their usage of causal statements.
Receptive language measure	The measure was developed by the researchers using a subset of vocabulary words presented in the books used during the intervention, and children were asked to identify the picture that corresponded to the target word (as cited in Wasik & Bond, 2001).
Expressive language measure	The measure was developed by the researchers and contained pictures representing the words in the books used during the intervention. Children were shown the pictures and asked to name the objects (as cited in Wasik & Bond, 2001).

Appendix A2.2 Outcome measure in the print knowledge domain

Outcome measure	Description
Print concepts factor	A factor derived from a number of outcome measures (subscales from the Developing Skills Checklist, ITPA-VE, PPVT-R, and EOWPVT-R) using a principal components analysis. Factor loadings for the print concepts factor were high for Name Letters, Blend CVC Words, Rhyming, Identify People Reading, Distinguish Words-Pictures-Numbers, Identify Functions of Words-Numbers, and Identify Components of Writing (as cited in Whitehurst, Epstein, et al., 1994).

Appendix A2.3 Outcome measures in the phonological processing domain

Outcome measure	Description
Rhyme oddity detection	A researcher-developed measure designed to measure children's understanding of words that rhyme (as cited in Lonigan et al., 1999).
Alliteration oddity detection	A researcher-developed measure designed to measure children's understanding of differences at the beginnings of words (as cited in Lonigan et al., 1999).
Sound blending	A researcher-developed measure designed to measure children's ability to combine word elements to form a new word (as cited in Lonigan et al., 1999).
Sound elision	A researcher-developed measure designed to measure children's ability to take away parts of words and say the word that is left over (as cited in Lonigan et al., 1999).
Linguistic awareness factor	A factor extracted from a number of outcome measures (subscales from the Developing Skills Checklist, ITPA-VE, PPVT-R, and EOWPVT-R) using a principal components analysis. Factor loadings for the linguistic awareness factor were high on Identify Sounds and Letters, Identify Same-Different Sounds, Segment Sentences, Segment Words (as cited in Whitehurst, Epstein, et al., 1994).

Appendix A2.4 Outcome measure in the early reading/writing domain

Outcome measure	Description
Writing factor	A factor derived from a number of outcome measures (subscales from the Developing Skills Checklist, ITPA-VE, PPVT-R, and EOWPVT-R) using a principal components analysis. Factor loadings for the writing factor were high for Print in Left-Right Progression, Print First Name, and Write Message Mechanics (as cited in Whitehurst, Epstein, et al., 1994).

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

Outcome measure	Study sample	Sample size (classrooms/ children)	Author's findings from the study: Mean outcome (standard deviation[2]) *Dialogic Reading* group[3]	Author's findings from the study: Mean outcome (standard deviation[2]) Comparison group[3]	WWC calculations: Mean difference[4] (*Dialogic Reading – comparison)*	WWC calculations: Effect size[5]	WWC calculations: Statistical significance[6] (at $\alpha = 0.05$)	WWC calculations: Improvement index[7]
Lonigan et al., 1999 (randomized controlled trial)[8]								
PPVT-R	2–5-year-olds	66	84.40 (17.28)	85.19 (14.01)	-0.79	-0.05	ns	-2
EOWPVT-R	2–5-year-olds	66	88.51 (10.57)	87.97 (15.11)	0.54	0.04	ns	+2
ITPA-VE	2–5-year-olds	66	45.46 (8.27)	40.81 (10.95)	4.65	0.48	Statistically significant	+18
WJ-LC	2–5-year-olds	29	8.51 (3.84)	7.29 (4.27)	1.22	0.29	ns	+11
Average[9] for oral language (Lonigan et al., 1999)						0.19	ns	+8
Lonigan & Whitehurst, 1998 (randomized controlled trial)[10]								
PPVT-R-Form M	3–4-year-olds	75	80.95 (10.54)	81.80 (13.35)	-0.85	-0.07	ns	-3
EOWPVT-R	3–4-year-olds	75	87.37 (9.69)	86.92 (11.19)	0.45	0.04	ns	+2
ITPA-VE	3–4-year-olds	75	109.09 (16.01)	102.60 (12.25)	6.49	0.43	ns	+17
Average[9] for oral language (Lonigan & Whitehurst, 1998)						0.14	ns	+5

(continued)

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

Outcome measure	Study sample	Sample size (classrooms/ children)	Author's findings from the study: Mean outcome (standard deviation[2]) — *Dialogic Reading* group[3]	Author's findings from the study: Mean outcome (standard deviation[2]) — Comparison group[3]	WWC calculations: Mean difference[4] (*Dialogic Reading – comparison)*	WWC calculations: Effect size[5]	WWC calculations: Statistical significance[6] (at $\alpha = 0.05$)	WWC calculations: Improvement index[7]
Wasik & Bond, 2001 (randomized controlled trial)[11]								
Receptive language	3–4-year-olds	4/121[12]	nr	nr	nr	1.58	Statistically significant	+44
Expressive language	3–4-year-olds	4/121[12]	nr	nr	nr	2.05	Statistically significant	+48
Average[9] for oral language (Wasik & Bond, 2001)						1.81	Statistically significant	+47
Whitehurst, Epstein, et al., 1994 and Zevenbergen et al., 2003 (randomized controlled trials)[13]								
Language factor	4-year-olds	15/167	-0.02 (1.00)	-0.10 (1.00)	0.08	0.08	ns	+3
Character states	4-year-olds	16/123	1.42 (1.82)	0.67 (0.86)	0.75	0.50	ns	+19
Dialogue	4-year-olds	16/123	1.56 (1.44)	0.96 (0.92)	0.60	0.48	ns	+18
Causal state	4-year-olds	16/123	0.18 (0.41)	0.33 (0.58)	-0.15	-0.30	ns	-12
Information/general content	4-year-olds	16/123	87.54 (14.32)	87.40 (11.50)	0.14	0.01	ns	0
Average[9] for oral language (Whitehurst, Epstein, et al., 1994; Zevenbergen et al., 2003)						0.15	ns	+6

(continued)

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

Outcome measure	Study sample	Sample size (classrooms/ children)	Author's findings from the study: Mean outcome (standard deviation[2]) — *Dialogic Reading* group[3]	Author's findings from the study: Mean outcome (standard deviation[2]) — Comparison group[3]	WWC calculations: Mean difference[4] (*Dialogic Reading – comparison)*	WWC calculations: Effect size[5]	WWC calculations: Statistical significance[6] (at $\alpha = 0.05$)	WWC calculations: Improvement index[7]
Crain-Thoreson & Dale, 1999 (randomized controlled trial with attrition problems)[14]								
Number of child utterances	3–5-year-olds	22	43.36 (22.10)	36.43 (21.30)	6.93	0.31	ns	+12
Ratio of child participation	3–5-year-olds	22	0.30 (0.11)	0.26 (0.12)	0.05	0.40	ns	+15
Mean length of utterances	3–5-year-olds	22	3.00 (0.93)	2.73 (0.97)	0.27	0.27	ns	+11
Lexical diversity	3–5-year-olds	22	55.29 (23.20)	52.56 (38.70)	2.73	0.09	ns	+3
PPVT-R	3–5-year-olds	22	63.70 (21.80)	59.54 (16.20)	4.16	0.20	ns	+8
EOWPVT-R	3–5-year-olds	22	70.12 (11.30)	71.73 (10.60)	-1.61	-0.14	ns	-6
Average[9] for oral language (Crain-Thoreson & Dale, 1999)						0.19	ns	+7
Whitehurst, Arnold, et al., 1994 (randomized controlled trial with attrition problems)[15]								
EOWPVT-R	3-year-olds	67	89.89 (13.40)	85.18 (16.73)	4.71	0.32	Statistically significant	+13
PPVT-R	3-year-olds	67	86.49 (13.81)	83.68 (15.83)	2.81	0.19	ns	+8
Our Word	3-year-olds	67	10.18 (5.49)	8.91 (7.00)	1.27	0.21	ns	+8
ITPA-VE	3-year-olds	67	100.06 (15.60)	100.11 (16.98)	-0.05	0.00	ns	0
Average[9] for oral language (Whitehurst, Arnold, et al., 1994)						0.18	ns	+7
Domain average[9] for oral language across all studies						0.44	na	+17

(continued)

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

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

1. This appendix reports findings considered for the effectiveness rating and the average improvement indices. Subgroup and follow-up findings from the same studies are not included in these ratings, but are reported in Appendix A4.1.
2. The standard deviation across all students in each group shows how dispersed the participants' outcomes are: a smaller standard deviation on a given measure would indicate that participants had more similar outcomes. For Lonigan and Whitehurst (1998) and Whitehurst, Arnold, et al. (1994) the intervention group standard deviations were calculated by the WWC based on subgroup standard deviations.
3. For Lonigan and Whitehurst (1998), the means reported in the table reflect those from high and low compliance centers combined. For Crain-Thoreson and Dale (1999) and Zevenbergen et al. (2003), the posttest means are covariate-adjusted means provided by the study authors. For Whitehurst, Epstein, et al. (1994), intervention and comparison group means were computed by the WWC based on graphs in the article. For Wasik and Bond (2001), the effect sizes were calculated based on the sample sizes and F (1,120) = 76.61 and F (1,120) = 128.43. For Lonigan and Whitehurst (1998) and Whitehurst, Arnold, et al. (1994), the intervention group means were calculated by the WWC by combining the school and school plus home conditions. For Lonigan et al. (1999), the intervention group mean equals the comparison group mean plus the mean difference.
4. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group. For Lonigan et al. (1999), the mean differences were computed by the WWC and took into account the pretest difference between the study groups. The resulting effect sizes may overestimate the intervention's effects when the intervention group had lower pretest scores than the comparison group, and underestimate the intervention's effect when the intervention group had higher pretest scores than 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 favorable results.
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. See Technical Details of WWC-Conducted Computations for the formulas the WWC used to calculate statistical significance. In the case of Lonigan et al. (1999), a correction for multiple comparisons was needed, but the significance levels do not differ from those reported in the original study.
9. 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 size.
10. In the case of Lonigan and Whitehurst (1998), a correction for multiple comparisons was needed, so the significance levels may differ from those reported in the original study. The findings in the table represent a comparison of children in the combined *Dialogic Reading* at school and at school plus home group versus children in the comparison group in both high and low compliance centers.
11. In the case of Wasik and Bond (2001), corrections for clustering and multiple comparisons were needed, but the significance levels do not differ from those reported in the original study.
12. For Wasik and Bond (2001), four teachers were randomly assigned to intervention and comparison groups.
13. In the case of Whitehurst, Epstein, et al. (1994) and Zevenbergen et al. (2003), corrections for clustering were needed, but the significance levels do not differ from those reported in the original studies.
14. In the case of Crain-Thoreson and Dale (1999), no corrections for clustering or multiple comparisons were needed.
15. In the case of Whitehurst, Arnold, et al. (1994), a correction for multiple comparisons was needed, so the significance levels may differ from those reported in the original study.

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

Outcome measure	Study sample	Sample size (classrooms/ children)	Author's findings from the study: Mean outcome (standard deviation[2]) — *Dialogic Reading* group[3]	Author's findings from the study: Mean outcome (standard deviation[2]) — Comparison group[3]	WWC calculations: Mean difference[4] (*Dialogic Reading* – comparison)	WWC calculations: Effect size[5]	WWC calculations: Statistical significance[6] (at $\alpha = 0.05$)	WWC calculations: Improvement index[7]
Whitehurst, Epstein, et al., 1994 (randomized controlled trial)[8]								
Print concepts factor	4-year-olds	15/167	0.26 (1.00)	-0.38 (1.00)	0.64	0.64	Statistically significant	+24
Domain average[9] for print knowledge						0.64	Statistically significant	+24

1. This appendix reports findings considered for the effectiveness rating and the average improvement index.
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. The intervention and comparison group means were computed by the WWC based on graphs in the article.
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 favorable results.
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. See Technical Details of WWC-Conducted Computations for the formulas the WWC used to calculate statistical significance. In the case of Whitehurst, Epstein, et al. (1994), a correction for clustering was needed, but the significance level does not differ from the one reported in the original study.
9. This row provides the study average, which in this case 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 (classrooms/ children)	Author's findings from the study: Mean outcome (standard deviation[2]) — *Dialogic Reading* group[3]	Author's findings from the study: Mean outcome (standard deviation[2]) — Comparison group[3]	WWC calculations: Mean difference[4] (*Dialogic Reading* – comparison)	WWC calculations: Effect size[5]	WWC calculations: Statistical significance[6] (at $\alpha = 0.05$)	WWC calculations: Improvement index[7]
Lonigan et al., 1999 (randomized controlled trial)[8]								
Rhyme oddity detection	2–5-year-olds	61	3.74 (1.40)	3.90 (1.42)	-0.16	-0.11	ns	-4
Alliteration oddity detection	2–5-year-olds	61	3.93 (1.30)	2.28 (1.28)	1.65	1.26	ns[9]	+40
Sound blending	2–5-year-olds	61	2.37 (6.03)	2.83 (5.27)	-0.46	-0.08	ns	-3
Sound elision	2–5-year-olds	61	2.85 (3.48)	3.55 (4.61)	-0.70	-0.17	ns	-7
Average[10] for phonological processing (Lonigan et al., 1999)						0.22	ns	+9
Whitehurst, Epstein, et al., 1994 (randomized controlled trial)[11]								
Linguistic awareness factor	4-year-olds	15/167	0.08 (1.00)	0.06 (1.00)	0.02	0.02	ns	+1
Average[10] for phonological processing (Whitehurst, Epstein, et al., 1994)						0.02	ns	+1
Domain average[10] for phonological processing across all studies						0.12	na	+5

ns = not statistically significant
na = not applicable

1. This appendix reports findings considered for the effectiveness rating and the average improvement indices.
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. For Lonigan et al. (1999), the intervention group mean equals the comparison group mean plus the mean difference. For Whitehurst, Epstein, et al. (1994), the intervention and comparison group means were computed by the WWC based on graphs in the article.
4. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group. For Lonigan et al. (1999), the mean differences were computed by the WWC and took into account the pretest difference between the study groups. The resulting effect sizes may overestimate the intervention's effects when the intervention group had lower pretest scores than the comparison group, and underestimate the intervention's effect when the intervention group had higher pretest scores than 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.

(continued)

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

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 favorable results.
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. See Technical Details of WWC-Conducted Computations for the formulas the WWC used to calculate statistical significance. In the case of Lonigan et al. (1999), no corrections for clustering or multiple comparisons were needed.
9. This statistical significance level was reported by the study authors. It differs from, but is more accurate than, the significance level based on the effect size presented in this table, which is an approximate effect estimate computed by the WWC.
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 size.
11. In the case of Whitehurst, Epstein, et al. (1994), a correction for clustering was needed, but the significance level does not differ from the one reported in the original study.

Appendix A3.4 Summary of study findings included in the rating for the early reading/writing domain[1]

Outcome measure	Study sample	Sample size (classrooms/ children)	Author's findings from the study: Mean outcome (standard deviation[2]) — *Dialogic Reading* group[3]	Author's findings from the study: Mean outcome (standard deviation[2]) — Comparison group[3]	WWC calculations: Mean difference[4] (*Dialogic Reading* – comparison)	WWC calculations: Effect size[5]	WWC calculations: Statistical significance[6] (at $\alpha = 0.05$)	WWC calculations: Improvement index[7]
Whitehurst, Epstein, et al., 1994 (randomized controlled trial)[8]								
Writing factor	4-year-olds	15/167	0.20 (1.00)	-0.34 (1.00)	0.54	0.54	Statistically significant	+20
Domain average[9] for early reading/writing						0.54	Statistically significant	+20

1. This appendix reports findings considered for the effectiveness rating and the average improvement index.
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. The intervention and comparison group means were computed by the WWC based on graphs in the article.
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 favorable results.
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. See Technical Details of WWC-Conducted Computations for the formulas the WWC used to calculate statistical significance. In the case of Whitehurst, Epstein, et al. (1994), a correction for clustering was needed, but the significance level does not differ from the one reported in the original study.
9. This row provides the study average, which in this case 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 subgroup findings for the oral language domain[1]

Outcome measure	Study sample	Sample size (classrooms/ children)	Author's findings from the study: Mean outcome (standard deviation[2]) *Dialogic Reading* group	Author's findings from the study: Mean outcome (standard deviation[2]) Comparison group	WWC calculations: Mean difference[3] (*Dialogic Reading* – comparison)	WWC calculations: Effect size[4]	WWC calculations: Statistical significance[5] (at $\alpha = 0.05$)	WWC calculations: Improvement index[6]
Lonigan & Whitehurst, 1998 (randomized controlled trial; high-compliance centers—*Dialogic Reading* at school)[7]								
PPVT-R-Form M	3–4-year-olds	31	80.80 (8.88)	80.70 (17.78)	0.10	0.01	ns	+0
EOWPVT-R	3–4-year-olds	31	91.20 (8.25)	86.80 (14.02)	4.40	0.37	ns	+14
ITPA-VE	3–4-year-olds	31	106.80 (12.74)	100.20 (12.21)	6.60	0.52	ns	+20
Lonigan & Whitehurst, 1998 (randomized controlled trial; high-compliance centers—*Dialogic Reading* both at school and at home)[8]								
PPVT-R-Form M	3–4-year-olds	23	79.00 (8.63)	80.70 (17.78)	-1.70	-0.10	ns	-4
EOWPVT-R	3–4-year-olds	23	91.30 (7.00)	86.80 (14.02)	4.50	0.35	Statistically significant	+14
ITPA-VE	3–4-year-olds	23	121.80 (16.69)	100.20 (12.21)	21.60	1.53	Statistically significant	+44
Lonigan & Whitehurst, 1998 (randomized controlled trial; high compliance centers—*Dialogic Reading* at school and both at school and at home)[9]								
PPVT-R-Form M	3–4-year-olds	38	80.23 (8.64)	80.70 (17.78)	-0.47	-0.03	ns	-1
EOWPVT-R	3–4-year-olds	38	91.23 (7.71)	86.80 (14.02)	4.43	0.40	ns	+16
ITPA-VE	3–4-year-olds	38	111.57 (15.46)	100.20 (12.21)	11.37	0.78	Statistically significant	+28
Lonigan & Whitehurst, 1998 (randomized controlled trial; low compliance centers—*Dialogic Reading* at school)[10]								
PPVT-R-Form M	3–4-year-olds	27	80.10 (12.62)	83.40 (7.47)	-3.30	-0.29	ns	-12
EOWPVT-R	3–4-year-olds	27	81.30 (11.59)	87.10 (8.97)	-5.80	-0.53	Statistically significant	-20
ITPA-VE	3–4-year-olds	27	102.30 (15.91)	106.10 (14.28)	-3.80	-0.24	ns	-10

(continued)

Appendix A4.1 Summary of subgroup findings for the oral language domain[1] (continued)

Outcome measure	Study sample	Sample size (classrooms/ children)	Author's findings from the study: Mean outcome (standard deviation[2]) — *Dialogic Reading* group	Author's findings from the study: Mean outcome (standard deviation[2]) — Comparison group	WWC calculations: Mean difference[3] (*Dialogic Reading* – comparison)	WWC calculations: Effect size[4]	WWC calculations: Statistical significance[5] (at $\alpha = 0.05$)	WWC calculations: Improvement index[6]
Lonigan & Whitehurst, 1998 (randomized controlled trial; low compliance centers—*Dialogic Reading* both at school and at home)[11]								
PPVT-R-Form M	3–4-year-olds	21	83.90 (11.33)	83.40 (7.47)	0.50	0.05	ns	+2
EOWPVT-R	3–4-year-olds	21	88.60 (4.88)	87.10 (8.97)	1.50	0.20	ns	+8
ITPA-VE	3–4-year-olds	21	114.50 (15.13)	106.10 (14.28)	8.40	0.55	ns	+21
Lonigan & Whitehurst, 1998 (randomized controlled trial; low compliance centers—*Dialogic Reading* at school and both at school and at home)[12]								
PPVT-R-Form M	3–4-year-olds	37	81.56 (12.06)	83.40 (7.47)	-1.84	-0.16	ns	-7
EOWPVT-R	3–4-year-olds	37	84.11 (10.11)	87.10 (8.97)	-2.99	-0.30	ns	-12
ITPA-VE	3–4-year-olds	37	106.99 (16.46)	106.10 (14.28)	0.89	0.06	ns	+2
Whitehurst, Arnold, et al., 1994 (randomized controlled trial with attrition problems; *Dialogic Reading* at school)[13]								
EOWPVT-R	3-year-olds	48	88.12 (10.43)	85.18 (16.73)	2.94	0.21	ns	+8
PPVT-R	3-year-olds	48	85.73 (11.54)	83.68 (15.83)	2.05	0.15	ns	+6
Our Word	3-year-olds	48	9.35 (5.00)	8.91 (7.00)	0.44	0.07	ns	+3
ITPA-VE	3-year-olds	48	99.62 (14.50)	100.11 (16.98)	-0.49	-0.03	ns	-1

(continued)

Appendix A4.1 Summary of subgroup findings for the oral language domain[1] (continued)

Outcome measure	Study sample	Sample size (classrooms/ children)	Author's findings from the study: Mean outcome (standard deviation[2]) — *Dialogic Reading* group	Author's findings from the study: Mean outcome (standard deviation[2]) — Comparison group	WWC calculations: Mean difference[3] (*Dialogic Reading* – comparison)	WWC calculations: Effect size[4]	WWC calculations: Statistical significance[5] (at $\alpha = 0.05$)	WWC calculations: Improvement index[6]
Whitehurst, Arnold, et al., 1994 (randomized controlled trial with attrition problems; *Dialogic Reading* both at school and at home)[14]								
EOWPVT-R	3-year-olds	41	92.32 (16.65)	85.18 (16.73)	7.14	0.42	ns	+16
PPVT-R	3-year-olds	41	87.53 (16.72)	83.68 (15.83)	3.85	0.23	ns	+9
Our Word	3-year-olds	41	11.32 (6.05)	8.91 (7.00)	2.41	0.36	ns	+14
ITPA-VE	3-year-olds	41	100.66 (17.38)	100.11 (16.98)	0.55	0.03	ns	+1

ns = not statistically significant

1. This appendix presents subgroup findings for measures that fall in oral language. Total group scores were used for rating purposes and are presented in Appendix A3.1.
2. The standard deviation across all students in each group shows how dispersed the participants' outcomes are: a smaller standard deviation on a given measure would indicate that participants had more similar outcomes.
3. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group.
4. For an explanation of the effect size calculation, see Technical Details of WWC-Conducted Computations.
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 favorable results.
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. See Technical Details of WWC-Conducted Computations for the formulas the WWC used to calculate statistical significance. In the case of Lonigan and Whitehurst (1998) (high-compliance centers—*Dialogic Reading* at school), no correction for clustering was needed.
8. In the case of Lonigan and Whitehurst (1998) (high-compliance centers—*Dialogic Reading* both at school and at home), no correction for clustering was needed.
9. In the case of Lonigan and Whitehurst (1998) (high compliance centers—*Dialogic Reading* at school and both at school and at home), no correction for clustering was needed.
10. In the case of Lonigan and Whitehurst (1998) (low compliance centers—*Dialogic Reading* at school), no correction for clustering was needed.
11. In the case of Lonigan and Whitehurst (1998) (low compliance centers—*Dialogic Reading* both at school and at home), no correction for clustering was needed.
12. In the case of Lonigan and Whitehurst (1998) (low compliance centers—*Dialogic Reading* at school and both at school and at home), no correction for clustering was needed.
13. In the case of Whitehurst, Arnold, et al. (1994) (*Dialogic Reading* at school), no correction for clustering was needed.
14. In the case of Whitehurst, Arnold, et al. (1994) (*Dialogic Reading* both at school and at home), no correction for clustering was needed.

Appendix A4.2 Summary of follow-up findings for the oral language domain[1]

Outcome measure	Study sample	Sample size (classrooms/ children)	Author's findings from the study: Mean outcome (standard deviation[2]) — *Dialogic Reading* group[3]	Author's findings from the study: Mean outcome (standard deviation[2]) — Comparison group[3]	WWC calculations: Mean difference[4] (*Dialogic Reading* – comparison)	WWC calculations: Effect size[5]	WWC calculations: Statistical significance[6] (at $\alpha = 0.05$)	WWC calculations: Improvement index[7]
Whitehurst, Arnold, et al., 1994 (randomized controlled trial with attrition problems; *Dialogic Reading* at school and both at school and at home)[8]								
EOWPVT-R	3-year-olds	52	91.16 (11.79)	88.07 (17.49)	3.09	0.23	Statistically significant	+9
PPVT-R	3-year-olds	53	81.07 (14.83)	83.21 (17.63)	-2.14	-0.13	ns	-5
ITPA-VE	3-year-olds	52	105.13 (16.27)	104.23 (24.95)	0.90	0.05	ns	+2
Whitehurst, Arnold, et al., 1994 (randomized controlled trial with attrition problems; *Dialogic Reading* at school)[9]								
EOWPVT-R	3-year-olds	37	91.17 (10.36)	88.07 (17.49)	3.10	0.23	ns	+9
PPVT-R	3-year-olds	37	79.52 (14.99)	83.21 (17.63)	-3.69	-0.23	ns	-9
ITPA-VE	3-year-olds	36	102.28 (15.67)	104.23 (24.95)	-1.95	-0.10	ns	-4
Whitehurst, Arnold, et al., 1994 (randomized controlled trial with attrition problems; *Dialogic Reading* both at school and at home)[10]								
EOWPVT-R	3-year-olds	29	91.14 (13.74)	88.07 (17.49)	3.07	0.19	ns	+8
PPVT-R	3-year-olds	30	83.31 (14.59)	83.21 (17.63)	0.10	0.01	ns	+0
ITPA-VE	3-year-olds	29	109.22 (17.12)	104.23 (24.95)	4.99	0.23	ns	+9

ns = not statistically significant

1. This appendix presents six-month follow-up findings for combined groups (*Dialogic Reading* at school and *Dialogic Reading* both at school and at home) and for subgroups (*Dialogic Reading* at school; *Dialogic Reading* both at school and at home) for measures that fall in oral language. Immediate posttest scores for the combined group were used for rating purposes and are presented in Appendix A3.1.

(continued)

Appendix A4.2 Summary of follow-up findings for the oral language domain[1] (continued)

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. For Whitehurst, Arnold, et al.(1994), the standard deviations for the *Dialogic Reading* at school and both at school and at home group were calculated by the WWC based on standard deviations of the *Dialogic Reading* at school group and the *Dialogic Reading* both at school and at home group.
3. For Whitehurst, Arnold, et al.(1994), the means for the *Dialogic Reading* at school and both at school and at home group were calculated by the WWC based on means of the *Dialogic Reading* at school group and the *Dialogic Reading* both at school and at home group.
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 favorable results.
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 (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. See Technical Details of WWC-Conducted Computations for the formulas the WWC used to calculate statistical significance. In the case of Whitehurst, Arnold, et al. (1994) (*Dialogic Reading* at school and both at school and at home), no correction for clustering was needed.
9. In the case of Whitehurst, Arnold, et al. (1994) (*Dialogic Reading* at school), no correction for clustering was needed.
10. In the case of Whitehurst, Arnold, et al. (1994) (*Dialogic Reading* both at school and at home), no correction for clustering was needed.

Appendix A5.1 *Dialogic Reading* rating for the oral language domain

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.[1]

For the outcome domain of oral language, the WWC rated *Dialogic Reading* as having positive effects. The remaining ratings (potentially positive effects, mixed effects, no discernible effects, potentially negative effects, and negative effects) were not considered because *Dialogic Reading* was assigned the highest applicable rating.

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

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

  Met. Three of the six studies that reported oral language outcomes found statistically significant positive effects. Two of these three studies met WWC evidence standards for a strong design.

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

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

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 effects for ratings of potentially positive or potentially negative effects. See the WWC Intervention Rating Scheme for a complete description.

Appendix A5.2 *Dialogic Reading* rating for the print knowledge domain

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.[1]

For the outcome domain of print knowledge, the WWC rated *Dialogic Reading* as having potentially positive effects. It did not meet the criteria for positive effects because it had only one study. The remaining ratings (mixed effects, no discernible effects, potentially negative effects, and negative effects) were not considered because *Dialogic Reading* was assigned the highest applicable rating.

Rating received

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.

  Met. One study examined print knowledge and produced a positive effect that was statistically significant based on both the authors' report and the WWC analysis.

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

  Met. No studies had statistically significant or substantively important negative effects or indeterminate effects in this domain.

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. Only one study assessed print knowledge.

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

  Met. No studies had statistically significant or substantively important negative effects in this domain.

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 effects for ratings of potentially positive or potentially negative effects. See the WWC Intervention Rating Scheme for a complete description.

Appendix A5.3 *Dialogic Reading* rating for the phonological processing domain

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.[1]

For the outcome domain of phonological processing, the WWC rated *Dialogic Reading* as having no discernible effects. It did not meet the criteria for positive effects, potentially positive effects, mixed effects, potentially negative effects, or negative effects because no studies showed statistically significant or substantively important effects, either positive or negative.

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 studies 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. Neither of the two studies meeting WWC evidence standards for a strong design showed statistically significantly positive effects.

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

  Met. No studies had statistically significant or substantively important negative effects in this domain.

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. No studies showed statistically significant or substantively important positive effects in this domain.

- 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. No studies had statistically significant or substantively important negative effects. Two studies showed indeterminate 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. No studies showed statistically significant or substantively important effects, either positive or negative, in this domain.

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. No studies had statistically significant or substantively important effects. Two studies showed indeterminate effects.

(continued)

Appendix A5.3 *Dialogic Reading* rating for the phonological processing domain (continued)

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

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

  Not met. No studies showed statistically significant or substantively important negative effects in this domain.

- Criterion 2: No studies showing a statistically significant or substantively important *positive* effect, or more studies showing statistically significant or substantively important *negative* effects than showing statistically significant or substantively important *positive* effects.

  Met. No studies showed statistically significant or substantively important positive effects in this domain.

Negative effects: Strong evidence of a negative effect with no overriding contrary evidence.

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

  Not met. No studies showed statistically significant negative effects in this domain.

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

  Met. No studies showed statistically significant or substantively important and positive effects in this domain.

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 effects for ratings of potentially positive or potentially negative effects. See the WWC Intervention Rating Scheme for a complete description.

Appendix A5.4 *Dialogic Reading* rating for the early reading/writing domain

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.[1]

For the outcome domain of early reading/writing, the WWC rated *Dialogic Reading* as having potentially positive effects. It did not meet the criteria for positive effects because it only had one study. The remaining ratings (mixed effects, no discernible effects, potentially negative effects, and negative effects) were not considered because *Dialogic Reading* was assigned the highest applicable rating.

Rating received

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.

  Met. One study found a statistically significant positive effect and the WWC analysis confirmed the statistical significance of the finding.

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

  Met. No studies were identified as having statistically significant or substantively important negative effects or indeterminate effects.

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. Only one study assessed early reading/writing outcomes.

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

  Met. No studies were identified as having statistically significant or substantively important negative effects.

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