

What Works Clearinghouse



Shared Book Reading

Practice description *Shared Book Reading* is a general practice aimed at enhancing young children’s language and literacy skills and their appreciation of books. Typically, *Shared Book Reading* involves an adult reading a book to one child or a small group of children without

requiring extensive interactions from them. Two related practices are reviewed in the What Works Clearinghouse (WWC) intervention reports on *Dialogic Reading* and *Interactive Shared Book Reading*.

Research Three studies of *Shared Book Reading* met the WWC evidence standards.¹ These studies, which included a total of 124 pre-school children, examined intervention effects on children’s oral language and phonological processing. All children were from

low- to middle-income families and about half were female. This report focuses on immediate posttest findings to determine the effectiveness of the intervention.²

Effectiveness *Shared Book Reading* was found to have mixed effects on oral language and potentially positive effects on phonological processing.

	Oral language	Print knowledge	Phonological processing
Rating of effectiveness	Mixed effects	Not reported	Potentially positive effects
Improvement index³	Average: +3 percentile points Range: -19 to +15 percentile points	Not reported	Average: +17 percentile points Range: +1 to +43 percentile points
	Early reading/writing	Cognition	Math
Rating of effectiveness	Not reported	Not reported	Not reported
Improvement index³	Not reported	Not reported	Not reported

1. 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 aged 3–5 or in preschool.

2. The evidence presented in this report is based on available research. Findings and conclusions may change as new research becomes available.

3. 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 *Shared Book Reading* practices reviewed by the ECE team included one study on which Dr. Lonigan was a primary author. Dr. Lonigan's financial interests are not affected by the success or failure of *Shared Book Reading* practices, nor does he receive any royalties or other monetary return from the use of *Shared Book Reading* practices. Dr. Lonigan was not involved in the decision to include the study in the review, and he was not involved in the

coding, reconciliation, or discussion of the included study. Dr. Kisker led all review activities related to the study. The decision to review *Shared Book Reading* practices 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 *Shared Book 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

Shared Book Reading is a practice that does not have a developer responsible for providing information or materials.

Scope of use

Information is not available on the number or demographics of children or centers using this intervention.

Teaching

In center-based settings, *Shared Book Reading* practices can be used by teachers with individual children or in small- and

large-group settings. Variations of *Shared Book Reading* include an adult reading a story to a child or group of children, an adult reading a story to a child or group of children and giving them the opportunity to retell the story, or an adult reading the same story more than one time to a child or group of children.

Cost

Information is not available about the costs of teacher training and implementation of *Shared Book Reading* practices.

Research

Ten studies reviewed by the WWC investigated the effects of *Shared Book Reading* practices in center-based settings. Three of the studies (Irlen, 2003a, b⁴; Lonigan, Anthony, Bloomfield, Dyer, & Samwel, 1999) were randomized controlled trials that met WWC evidence standards. The remaining seven studies did not meet WWC evidence screens.

Irlen (2003a) included 33 four- to five-year-old children from three preschools in the Los Angeles area. Two groups were included: storybook-retell (adults read a book to a group of children and children retold the story) and video-retell (children

watched the story on a video and then retold the story). The study also included 10 children in a no-treatment comparison group that was added to the study sample after group assignment was complete and the intervention had started. For the purposes of this review, the WWC chose the storybook-retell condition as the intervention condition and the video-retell condition as the comparison condition and did not include the post-hoc no-treatment control group.⁵ In this study, oral language outcomes for children in the storybook-retell condition were compared with those of children in the video-retell condition.

4. Irlen (2003) included two different but relevant shared book reading conditions. The WWC counted the report by Irlen (2003) as two studies and designated the storybook retell vs. video retell comparison as Irlen (2003a) and the storybook repeat vs. video repeat comparison as Irlen (2003b). Use of the active contrast groups in Irlen (2003a, b) may have reduced the effect size estimate, but this contrast specifically isolates the impact of shared reading because it is the only factor that varies between the contrasted conditions.

5. This decision was made because the no-treatment comparison group was not part of the original randomized sample and the author did not provide evidence of group equivalence prior to the intervention for the no-treatment comparison group and the other study groups. The inclusion of this comparison group would have resulted in the study not meeting WWC evidence screens.

Research *(continued)*

Irlen (2003b) included 30 four- to five-year-old children from three preschools in the Los Angeles area. Two groups were included: storybook-repeat (children listened to the story twice) and video-repeat (children watched the story twice). The study also included 10 children in a no-treatment comparison group that was added to the study sample after group assignment was complete and the intervention had started. For the purposes of this review, the WWC chose the storybook-repeat condition as the intervention condition and the video-repeat condition as the comparison condition and did not include the post-hoc no-treatment control group.⁵ In this study, oral language outcomes

for children in the storybook-repeat condition were compared with those of children in the video-repeat condition.

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. This study 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 typical *Shared Book Reading* group and the no-treatment comparison group⁶ with a total of 61 children.

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.⁷

Oral language. Three studies examined outcomes in the oral language domain. One study showed statistically significant and positive effects, and two studies showed indeterminate effects. Irlen (2003a) compared the storybook-retell and video-retell groups on three measures in this outcome domain, and no statistically significant differences were found (as calculated by the WWC).⁶ In this study, the effect of *Shared Book Reading* on oral language was indeterminate, according to WWC criteria. Irlen (2003b) compared the storybook-repeat and video-repeat groups on the same three oral language measures. As with Irlen (2003a), the WWC found no statistically significant differences between the two groups.⁸ In this study, the effect of *Shared Book Reading* on oral language was indeterminate, according to WWC criteria. Lonigan et al. (1999)

reported findings for four measures in this outcome domain. The authors reported, and the WWC confirmed, a statistically significant difference favoring the intervention group on the Woodcock-Johnson Psychoeducational Battery-Listening Comprehension Subtest (WJ-LC). There were no statistically significant effects for the other three outcome measures. In this study, the effect of *Shared Book Reading* on oral language was statistically significantly and positive, according to WWC criteria.

Phonological processing. Lonigan et al. (1999) reported, and the WWC confirmed, a statistically significant difference favoring the intervention group on one of the four measures in this outcome domain (alliteration oddity detection). In this study, *Shared Book Reading* had a statistically significantly positive effect on phonological processing, 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

6. The comparison between the *Dialogic Reading* group and the comparison group is included in the WWC *Dialogic Reading* intervention report.

7. The level of statistical significance was reported by the study authors or, where necessary, calculated by the WWC to correct for clustering within classrooms or schools and for multiple comparisons. For an explanation about the clustering correction, see the [WWC Tutorial on Mismatch](#). See the [Technical Details of WWC-Conducted Computations](#) for the formulas the WWC used to calculate the statistical significance. In the case of the *Shared Book Reading* report, a correction for multiple comparisons was needed.

8. The study author did not report the statistical significance levels of the differences between these two groups, so the WWC computed the significance levels based on data provided by the author.

Effectiveness *(continued)*

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,⁷ the size of the

difference between participants in the intervention condition and the comparison condition, and the consistency in findings across studies (see the [WWC Intervention Rating Scheme](#)).

The WWC found *Shared Book Reading* to have mixed effects for oral language and potentially positive effects for phonological processing

Improvement index

The WWC computes an improvement index for each finding. In addition, within each outcome domain, the WWC computes an average improvement index for each study as well as an average improvement 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 +3 percentile points across three studies, with a range of -19 to +15 percentile points across findings. The average improvement index for phonological processing is +17 percentile points for one study, with a range of +1 to +43 percentile points across findings.

Summary

The WWC reviewed 10 studies on *Shared Book Reading*.⁹ Three of these studies met WWC evidence standards; the remaining studies did not meet WWC evidence screens. Based on these three studies, the WWC found mixed effects for oral language and potentially positive effects for phonological processing. The evidence presented in this report may change as new research emerges.

References

Met WWC evidence standards

- Irlen, S. M. (2003a). The impact of video viewing and retelling on preliterate children's narrative comprehension. *Dissertation Abstracts International*, 64(04), 1174A. (UMI No. 3088967). (Study: storybook-retell versus video-retell).
- Irlen, S. M. (2003b). The impact of video viewing and retelling on preliterate children's narrative comprehension. *Dissertation Abstracts International*, 64(04), 1174A. (UMI No. 3088967). (Study: storybook-repeat versus video-repeat).
- Lonigan, C. J., Anthony, J. L., Bloomfield, B. G., Dyer, S. M., & Samwel, C. S. (1999). Effects of two shared-reading interven-

tions on emergent literacy skills of at-risk preschoolers. *Journal of Early Intervention*, 22(4), 306-322.

Did not meet WWC evidence screens

Box, J. A., & Aldridge, J. (1993). Shared reading experiences and Head Start children's concepts about print and story structure. *Perceptual & Motor Skills*, 77(3), 929-930.¹⁰

Additional source:

Box, J. A. (1991). The effects of shared reading experiences on Head Start children's concepts about print and story structure. *Dissertation Abstracts International*, 51(12), 4015A. (UMI No. 9107738).¹⁰

⁹. One single-case design study was identified but is not included in this review because the WWC does not yet have standards for reviewing single-case design studies.

¹⁰. Complete data were not reported: the WWC could not compute effect sizes.

References

- Jones, C. B. (1987). The effect of oral reading by senior citizens on the oral language and readiness skills of language-delayed prekindergarten children. *Dissertation Abstracts International*, 48(3), 556A. (UMI No. 8713620).¹¹
- Lamb, H. A. (1986). The effects of a read-aloud program with language interaction. *Dissertations Abstracts International*, 47(5-A). (UMI No. 8616894).¹²
- Mautte, L. A. (1991). The effects of adult-interactive behaviors within the context of repeated storybook readings upon the language development and selected prereading skills of prekindergarten at-risk students. *Dissertation Abstracts International*, 52(1), 122A. (UMI No. 9115887).¹²
- Murray, B. A., Stahl, S. A., & Ivey, M. G. (1996). Developing phoneme awareness through alphabet books. *Reading & Writing*, 8(4), 307–322.¹³

Additional source:

- Murray, B. A., Stahl, S. A., & Ivey, M. G. (1993). Developing phonological awareness through alphabet books. Paper presented at the Annual Meeting of the National Reading Conference (Charleston, SC, December 1–4, 1993).¹³
- Neuman, S. B. (1996). Children engaging in storybook reading: The influence of access to print resources, opportunity, and parental interaction. *Early Childhood Research Quarterly*, 11(4), 495–514.¹⁴
- Ratliff, J. L. (1987). Explicit instruction in story structure: Effects on preschoolers' listening comprehension. *Dissertation Abstracts International*, 47(11), 3972A. (UMI No. 8625352).¹²

For more information about specific studies and WWC calculations, please see the [WWC Shared Book Reading Technical Appendices](#).

11. Lack of evidence for baseline equivalence: the study, which used a quasi-experimental design, did not establish that the comparison group was equivalent to the intervention group at the baseline.

12. Confound: the effects of the intervention could not be separated from other factors; the impact of the agent of the intervention was confounded with the impact of the intervention.

13. Confound: There was only one cluster in each study condition; therefore, the effects of the intervention could not be separated from the effects of the cluster.

14. Does not use a strong causal design: the study did not use a comparison group.

Appendix

Appendix A1.1 Study characteristics: Irlen, 2003a (randomized controlled trial)¹

Characteristic	Description
Study citation	Irlen, S. M. (2003a). The impact of video viewing and retelling on preliterate children's narrative comprehension. <i>Dissertation Abstracts International</i> , 64(04), 1174A. (UMI No. 3088967).
Participants	<p>Irlen (2003a, b). The complete sample reported by Irlen (2003a, b) included 73² four- to five-year-old children (mean age = 56.7 months) who were randomly assigned within gender to four conditions. All children were from lower to middle income homes; 52% were male, 39% were Caucasian, 24% were African American, 16% were Asian American, 10% were Latino, and 11% were of unknown race/ethnicity.</p> <p>Irlen (2003a). Results for the 33 children who had been randomly assigned to the storybook-retell and video-retell conditions are included in this report.</p>
Setting	The study took place in three preschools in the Los Angeles area (specifically, two preschools in Long Beach, California and one preschool in Gardena, California). Two of the schools served primarily lower income and minority students; the third school served primarily middle income and Caucasian children.
Intervention	The WWC designated the storybook-retell group as the intervention condition for this review. This group listened to the book as a group and retold the story as a group. The intervention took place approximately 35 minutes per day, for a period of 10 days (average exposure to the intervention was 8.5 days) in private rooms separate from the preschool classrooms in the three schools.
Comparison	The WWC designated the video-retell group as the comparison condition for this review. This group watched a video and then retold the story as a group. The intervention took place approximately 35 minutes per day, for a period of 10 days (average exposure to the intervention was 8.5 days) in private rooms separate from the preschool classrooms in the three schools.
Primary outcomes and measurement	The primary outcome domain was oral language, which was assessed with a Prompted score (Paris rubric score), a Prompted and Unaided score (Marshall checklist), and an Unaided Retelling score (Marshall checklist). These are all measures of children's narrative comprehension (see Appendix A2.1-2.2 for more detailed descriptions of outcome measures).
Teacher training	The intervention was implemented by the study author and two undergraduate assistants. The training provided to the undergraduate assistants was not described.

1. Irlen (2003) included two different but relevant shared book reading conditions. The WWC counted the report by Irlen (2003) as two studies and designated the storybook retell vs. video retell comparison as Irlen (2003a) and the storybook repeat vs. video repeat comparison as Irlen (2003b). Use of the active contrast groups in Irlen (2003a, b) may have reduced the effect size estimate, but this contrast specifically isolates the impact of shared reading because it is the only factor that varies between the contrasted conditions.
2. An additional 10 children were included in a post-hoc no-treatment comparison group. However, the no-treatment comparison group was not part of the original randomized sample and the author did not provide evidence of group equivalence prior to the intervention for the no-treatment comparison group and the other study groups. Consequently, we did not include this group in our review because its inclusion would have resulted in the study not meeting WWC evidence screens.

Appendix A1.2 Study characteristics: Irlen, 2003b (randomized controlled trial)³

Characteristic	Description
Study citation	Irlen, S. M. (2003b). The impact of video viewing and retelling on preliterate children's narrative comprehension. <i>Dissertation Abstracts International</i> , 64(04), 1174A. (UMI No. 3088967).
Participants	<u>Irlen (2003a, b)</u> . The complete sample reported by Irlen (2003a, b) included 73 ⁴ four- to five-year-old children (mean age = 56.7 months) who were randomly assigned within gender to four conditions. All children were from lower to middle income homes; 52% were male; 39% were Caucasian, 24% were African American, 16% were Asian American, 10% were Latino, and 11% were of unknown race/ethnicity. <u>Irlen (2003b)</u> . Results for the 30 children who had been randomly assigned to the storybook-repeat and video-repeat conditions are included in this report.
Setting	The study took place in three preschools in the Los Angeles area (specifically, two preschools in Long Beach, California and one preschool in Gardena, California). Two of the schools served primarily lower income and minority students; the third school served primarily middle income and Caucasian children.
Intervention	The WWC designated the storybook-repeat group as the intervention condition for this review. This group listened to the book twice in a row. The intervention took place approximately 35 minutes per day, for a period of 10 days (average exposure to the intervention was 8.5 days) in private rooms separate from the preschool classrooms in the three schools.
Comparison	The WWC designated the video-repeat group as the comparison condition for this review. This group watched a video twice in a row. The intervention took place approximately 35 minutes per day, for a period of 10 days (average exposure to the intervention was 8.5 days) in private rooms separate from the preschool classrooms in the three schools.
Primary outcomes and measurement	The primary outcome domain was oral language, which was assessed with a Prompted score (Paris rubric score), a Prompted and Unaided score (Marshall checklist), and an Unaided Retelling score (Marshall checklist). These are all measures of children's narrative comprehension (see Appendix A2.1-2.2 for more detailed descriptions of outcome measures).
Teacher training	The intervention was implemented by the study author and two undergraduate assistants. The training provided to the undergraduate assistants was not described.

3. Irlen (2003) included two different but relevant shared book reading conditions. The WWC counted the report by Irlen (2003) as two studies and designated the storybook retell vs. video retell comparison as Irlen (2003a) and the storybook repeat vs. video repeat comparison as Irlen (2003b). Use of the active contrast groups in Irlen (2003a, b) may have reduced the effect size estimate, but this contrast specifically isolates the impact of shared reading because it is the only factor that varies between the contrasted conditions.

4. An additional 10 children were included in a post-hoc no-treatment comparison group. However, the no-treatment comparison group was not part of the original randomized sample and the author did not provide evidence of group equivalence prior to the intervention for the no-treatment comparison group and the other study groups. Consequently, we did not include this group in our review because its inclusion would have resulted in the study not meeting WWC evidence screens.

Appendix A1.3 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. <i>Journal of Early Intervention, 22</i> (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 61 children who had been randomly assigned within center to the typical shared reading and no-treatment comparison conditions are included in this report. ⁵
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 received a state child care subsidy.
Intervention	The study included two intervention groups: typical shared reading and Dialogic Reading. The typical shared reading intervention is included in this review; results involving <i>Dialogic Reading</i> are included in the WWC <i>Dialogic Reading</i> intervention report. In the typical shared reading condition, trained undergraduate volunteers read books to groups of three to five children daily for 10 to 15 minutes per session in a location outside the classroom during a six-week period. The undergraduate volunteers read the text of the book to the children, commented on the pictures, and answered questions that children asked.
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 Appendix A2.1-2.2 for more detailed descriptions of outcome measures).
Teacher training	No information on teacher training was provided for the typical shared book reading condition.

5. The WJ-LC measure was only completed by the 31 children participating in the second wave of the study; however, Lonigan et al. (1999) reported that study wave did not enter into a significant interaction with treatment condition for any variable measured across both waves of the study.

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 Lonigan et al., 1999).
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 Lonigan et al., 1999).
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 et al., 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).
Prompted score (Paris rubric score)	This score was derived by using the coding scheme for the prompted question portion of the Paris and Paris (2003) Narrative Comprehension Task to assess children’s knowledge of story structure elements and narrative comprehension (as cited in Irlen, 2003a, b). It was based on children’s responses to prompted questions involving 10 story grammar elements.
Prompted and Unaided score (Marshall checklist)	The Marshall (1983) checklist assessed narrative comprehension based on unaided story retelling and children’s responses to prompted questions (as cited in Irlen, 2003a, b). This score took into account children’s aided and unaided abilities to mention 7 of the 10 story grammar elements. This score shows children’s knowledge of story structure elements and narrative comprehension.
Unaided Retelling score (Marshall checklist)	The Unaided Retelling score indicates the number of story grammar elements that children mentioned exclusively in the retelling portion of the Marshall checklist described above. It assesses children’s knowledge of story structure elements and narrative comprehension (as cited in Irlen, 2003a, b).

Appendix A2.2 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).

Appendix A3.1 Summary of study findings included in the rating for the oral language domain¹

Outcome measure	Study sample	Sample size (children)	Author's findings from the study		WWC calculations			
			Mean outcome (standard deviation ²)		Mean difference ⁴ (intervention – comparison)	Effect size ⁵	Statistical significance ⁶ (at $\alpha = 0.05$)	Improvement index ⁷
			Intervention group ³	Comparison group				
Irlen, 2003a (randomized controlled trial)⁸								
Prompted score	4–5 year olds	33	8.78 (4.54)	8.40 (3.60)	0.38	0.09	ns	+4
Prompted and Unaided score	4–5 year olds	30	6.77 (2.10)	6.83 (2.70)	–0.06	–0.02	ns	–1
Unaided retelling score	4–5 year olds	32	1.89 (1.56)	1.64 (1.09)	0.25	0.18	ns	+7
Average⁹ for oral language (Irlen, 2003a)						0.08	ns	+3
Irlen, 2003b (randomized controlled trial)¹⁰								
Prompted score	4–5 year olds	30	9.53 (3.71)	9.09 (4.11)	0.44	0.11	ns	+4
Prompted and Unaided score	4–5 year olds	29	6.82 (2.19)	6.55 (2.45)	0.27	0.11	ns	+5
Unaided Retelling score	4–5 year olds	29	1.63 (1.18)	1.72 (1.30)	–0.09	–0.07	ns	–3
Average⁹ for oral language (Irlen, 2003b)						0.05	ns	+2
Lonigan et al., 1999 (randomized controlled trial)¹¹								
PPVT-R	2–5 year olds	61	77.95 (15.07)	85.19 (14.01)	–7.24	–0.49	ns	–19
EOWPVT-R	2–5 year olds	61	88.78 (14.29)	87.97 (15.11)	0.81	0.05	ns	+2
ITPA-VE	2–5 year olds	61	43.94 (9.50)	40.81 (10.95)	3.13	0.30	ns	+12
WJ-LC	2–5 year olds	31	9.10 (5.02)	7.29 (4.27)	1.81	0.38	Statistically significant	+15
Average⁹ for oral language (Lonigan et al., 1999)						0.06	ns	+2
Domain average⁹ for oral language across all studies						0.06	na	+3

(continued)

Appendix A3.1 Summary of study findings included in the rating for the oral language domain¹ *(continued)*

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 Irlen (2003a, b), the posttest means are covariate-adjusted means provided by the study author. 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 groups.
7. The improvement index represents the difference between the percentile rank of the average student in the intervention condition and the percentile rank of the average student in the comparison condition. The improvement index can take on values between -50 and +50, with positive numbers denoting results favorable to the intervention group.
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 the Irlen (2003a) study, no corrections for clustering or multiple comparisons were needed.
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 the Irlen (2003b) study, no corrections for clustering or multiple comparisons were needed.
11. In the case of the Lonigan et al. (1999) study, 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 phonological processing domain¹

Outcome measure	Study sample	Sample size (children)	Author's findings from the study		WWC calculations			
			Mean outcome (standard deviation ²)		Mean difference ⁴ (intervention – comparison)	Effect size ⁵	Statistical significance ⁶ (at $\alpha = 0.05$)	Improvement index ⁷
			Intervention group ³	Comparison group				
Lonigan et al., 1999 (randomized controlled trial)⁸								
Rhyme oddity detection	2–5 year olds	56	4.10 (2.13)	3.90 (1.42)	0.20	0.11	ns	+4
Alliteration oddity detection	2–5 year olds	56	4.60 (1.83)	2.28 (1.28)	2.32	1.46	Statistically significant	+43
Sound blending	2–5 year olds	56	2.94 (6.53)	2.83 (5.27)	0.11	0.02	ns	+1
Sound elision	2–5 year olds	56	4.39 (5.20)	3.55 (4.61)	0.84	0.17	ns	+7
Domain average⁹ for phonological processing						0.44	ns	+17

ns = not statistically significant

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. 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. 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 groups.
7. The improvement index represents the difference between the percentile rank of the average student in the intervention condition and the percentile rank of the average student in the comparison condition. The improvement index can take on values between -50 and +50, with positive numbers denoting results favorable to the intervention group.
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 the Lonigan et al. (1999) study, a correction for multiple comparisons was needed, so the significance levels may differ from those reported in the original study.
9. This row provides the study average, which in this instance, is also the domain average. The WWC-computed domain average effect size is a simple average rounded to two decimal places. The domain improvement index is calculated from the average effect size.

Appendix A4.1 Shared Book Reading rating for the oral language domain

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

For the outcome domain of oral language, the WWC rated *Shared Book Reading* as having mixed effects. It did not meet the criteria for *positive effects* or *potentially positive effects* because only one study showed statistically significant and positive effects and two studies showed indeterminate effects. The remaining ratings (*no discernible effects*, *potentially negative effects*, *negative effects*) were not considered, as *Shared Book Reading* was assigned the highest applicable rating.

Rating received

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

- Criterion 1: At least one study showing a statistically significant or substantively important *positive* effect, and at least one study showing a statistically significant or substantively important *negative* effect, but no more such studies than the number showing a statistically significant or substantively important *positive* effect.

Not met. One study showed a statistically significant and positive effect, but no studies showed statistically significant and negative or substantively important and negative effects.

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.

Met. One study showed a statistically significant and positive effect, and two studies showed 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 showed statistically significant and positive effects.

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

Met. There were no studies identified as having statistically significant and negative or substantively important and negative effects.

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

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

Met. One study showed a statistically significant and positive effect.

- 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. None of the studies was identified as having statistically significant and negative or substantively important and negative effects. Two of the studies showed indeterminate 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 effects. See the [WWC Intervention Rating Scheme](#) for a complete description.

Appendix A4.2 Shared Book Reading rating for the phonological processing domain

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

For the outcome domain of phonological processing, the WWC rated *Shared Book Reading* as having potentially positive effects. It did not meet the criteria for *positive effects*, as it only had one study. The remaining ratings (*mixed effects*, *no discernible effects*, *potentially negative effects*, *negative effects*) were not considered, as *Shared Book 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 that included phonological processing showed statistically significant and positive effects.

- 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. The study that included phonological processing did not have statistically significant and negative, substantively important and negative, 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 included phonological processing.

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

Met. The study that included phonological processing did not have statistically significant and negative or substantively important and 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 effects. See the [WWC Intervention Rating Scheme](#) for a complete description.