WWC Intervention Report

INSTITUTE OF EDUCATION SCIENCES

Early Childhood Education Updated April 2015

Shared Book Reading

Program Description[1]

Shared book reading (also known as *interactive shared book reading*) encompasses practices that adults can use when reading with children, which are intended to enhance young children's language and literacy skills. During *shared book reading*, an adult reads a book to an individual child or a group of children and uses one or more planned or structured interactive techniques to actively engage the children in the text. The adult may direct the children's attention to illustrations, print, or word meanings. The adult may engage children in discussions focused on understanding the meaning or sequence of events in a story or on understanding an expository passage. Adults may ask children questions, give explanations and draw connections between events in the text and those in the children's own lives as a way of expanding on the text and scaffolding children's learning experiences to support language development, emergent reading, and comprehension. Importantly, the adult engages in one or more interactive techniques to draw attention to aspects of the text being read.

One specific type of *shared book reading, dialogic reading*, is addressed in a separate What Works Clearinghouse (WWC) intervention report.[2] During *dialogic reading*, the adult uses a specific approach to prompting children's participation and functions as an active listener and questioner, enabling the adult and the child to switch roles so that the child learns to become the storyteller.

Report Contents

This intervention report presents findings from a systematic review of *shared book reading* conducted using the WWC Procedures and Standards Handbook, version 3.0, and the Early Childhood Education review protocol, version 3.0.

Research[3]

The WWC identified eight studies of *shared book reading* that both fall within the scope of the Early Childhood Education topic area and meet WWC group design standards. Six studies meet WWC group design standards without reservations and two studies meet WWC group design standards with reservations. Together, these studies included 791 children aged 3–6 years in 10 locations.

The WWC considers the extent of evidence for the effects of *shared book reading* on the school readiness outcomes of preschool children to be small for three outcome domains—comprehension, alphabetics, and general reading achievement—and medium to large for one outcome domain—language development. There were no studies that meet standards in the four other domains,[4] so this intervention report does not report on the effectiveness of *shared book reading* for those domains. (See the Effectiveness Summary on p. 7 for more details of effectiveness by domain.)

Effectiveness

Shared book reading was found to have mixed effects on comprehension and language development and no discernible effects on alphabetics and general reading achievement for preschool children.

Table 1. Summary of findings[5]

Outcome domain	Rating of effectiveness	Improvement index *(percentile points)*		Number of studies	Number of students	Extent of evidence
		Average	Range			
Comprehension	Mixed effects	+7	–13 to +32	4	233	Small
Language development	Mixed effects	+3	–20 to +18	4	576	Medium to large
Alphabetics	No discernible effects	+5	–9 to +16	2	316	Small
General reading achievement	No discernible effects	–8	na	1	38	Small

na = not applicable

Program Information

Background

Shared book reading is a practice that does not have a single developer responsible for providing information or materials. The interventions described in this report were developed by more than one developer and are not available for distribution through a common developer. Readers interested in using *shared book reading* practices in their classrooms can refer to sources available through Internet searches for information.

Program details

In school or center-based settings, *shared book reading* can be used with individual children, small groups of children, or a whole class. The practice may supplement or supplant other classroom read-aloud practices or language and early literacy practices, depending on how it is implemented. Typically, an adult reads a book to a child or children and engages the child/children in the text through planned or structured interactive techniques before, during, and after reading the text. For instance, before reading, the adult may ask children to point to the title, make predictions about what might happen in the book, or draw connections to children's own lives. During book reading, the adult may ask questions, give explanations, pose prompts, or call on a child to answer a specific question. The focus of the questions, explanations, and prompts may vary depending on the skills being taught. The adult may encourage children's responses and spontaneous contributions to enhance their language development, or emphasize features of story structure to support children's narrative discourse and comprehension. The adult may focus on modeling reading to children and helping them with various aspects of print awareness, such as learning that text is read from top to bottom and left to right. Adults may focus on building vocabulary and general background knowledge, as well as developing comprehension strategies, including identifying the topic and main idea. Importantly, the interaction between the adult and children is focused on the text being read.

The books chosen for reading may be storybooks or expository texts. Books may be selected to relate to broader classroom themes. Books may have large print, a small number of words per page, and illustrations throughout; however, *shared book reading* can be practiced with a wide variety of books, including those without any text. Although *shared book reading* practices vary in frequency, reading sessions are generally brief (e.g., 10–15 minutes a day) and occur several days a week.

Cost

Information is not available about the costs of teacher training and implementation of *shared book reading* practices.

Research Summary

The WWC identified 13 eligible studies that investigated the effects of *shared book reading* on the school readiness of preschool children. An additional 42 studies were identified but do not meet WWC eligibility criteria for review in this topic area. Citations for all 55 studies are in the References section, which begins on p. 11.

Table 2. Scope of reviewed research

Grades	PK
Delivery method	Individual, Small group, Whole class
Program type	Practice

The WWC reviewed 13 of those studies against group design standards. Six studies (Bochna, 2010; Box & Aldridge, 1993; Lamb, 1986; Mason, Kerr, Sinha, & McCormick, 1990; Mautte, 1991; and Walsh, 2010) are randomized controlled trials that meet WWC group design standards without reservations, and two studies (Justice, McGinty, Piasta, Kaderavek, & Fan, 2010; Pollard-Durodola, et al., 2011) are randomized controlled trials that meet WWC group design standards with reservations. Those eight studies are summarized in this report. Five studies do not meet WWC group design standards.

Summary of studies meeting WWC group design standards without reservations

Bochna (2010) examined the effects of *shared book reading* on comprehension. The author conducted a randomized controlled trial with children aged 3–5 years, attending three full-day, full-year Head Start centers in Pennsylvania. Forty-five children were randomly assigned either to a *shared book reading* intervention condition (21 children) or to a comparison condition (24 children). Both conditions involved an adult reading with a small group of children. Within each condition, five small groups were formed for reading. The *shared book reading* intervention focused on using questions to teach children to identify the topic and main idea in expository texts. The comparison condition involved the adult reading the same expository books to children without questions or instruction in identifying the topic and main idea. The analytic sample included 36 children; 17 in the intervention group and 19 in the comparison group.

Box and Aldridge (1993) examined the effects of *shared book reading* on alphabetics and comprehension. The study took place in 15 classrooms located in Head Start centers or public schools serving low-income children in one geographic area. The 15 Head Start classrooms were randomly assigned to three conditions: the *shared book reading* intervention, a business-as-usual general classroom instruction comparison, and a movement activities comparison. Within each classroom, five 4-year-old children were randomly selected to participate in the study. The *shared book reading* intervention involved the classroom teacher reading with a group of five children for 10–15 minutes a day, 4 days per week, for 8 weeks. The teacher introduced new books and re-read them repeatedly during the week. The teacher focused on story structure and reading strategies and encouraged children to chime in on repetitive sections, fill in missing words, and predict possible story outcomes. One comparison condition was regular classroom instruction, during which children had access to the usual learning centers. A second comparison condition involved the teacher leading children in movement activities for 15 minutes a day, 4 days per week, for 8 weeks. The analytic sample included 72 children: 25 in the intervention group, 23 in the business-as-usual group, and 24 in the movement activities group.[6]

Lamb (1986) examined effects of *shared book reading* on alphabetics, language development, and comprehension outcomes. The study included 36 three- to five-year-old low-income children from a day care center in Tallahassee, Florida. Lamb compared three 10-week intervention groups—read-aloud with structured language interaction, read-aloud only, and language interaction only—to a no-intervention comparison group participating in regular preschool activities. This WWC intervention report focuses on the comparison of the read-aloud with structured language interaction group (the *shared book reading* condition, with 10 children) and two comparison conditions, the read-aloud only group (nine children) and the interaction without read-aloud group (nine children).[7] The comparison between the read-aloud with language interaction group and the no-intervention comparison group is not included in this intervention report because it did not meet WWC group design standards.[8]

Mason et al. (1990) examined the effects of *shared book reading* on language development.[9] The study was conducted in 12 classes in two Early Start schools, a half-day developmental program in Illinois for 4-year-old children at risk for school failure. The 12 classes were randomly assigned, half to a 28-week *shared book reading* intervention group and half to a standard literacy lessons and activities comparison group. In the *shared book reading* group, teachers read aloud with a small group of children and encouraged discussion of the book. The analytic sample included 232 four-year-old children.

Mautte (1991) conducted a randomized controlled trial that examined the effects of shared book reading on language development and general reading achievement. The study included 53 at-risk low-income 4-year-old children from an inner-city early childhood education center in Tampa, Florida. Mautte compared two intervention groups—repeated reading with adult interaction and repeated reading without adult interaction—to a no-intervention comparison group participating in regular preschool curriculum activities. This WWC intervention report focuses on the comparison between the repeated reading with adult interaction group (the *shared book reading* condition) and the repeated reading without adult interaction group (the comparison condition) with a total of 38 children. The comparison between the repeated reading with adult interaction group and the no-intervention comparison group was not used by the WWC because it did not meet WWC group design standards.[10]

Walsh (2010) used a three-group study to examine the effect of *shared book reading* on comprehension. Two of the groups were one-to-one *shared book reading* interventions: one used questioning practices that were vocabulary-eliciting and one used questioning practices that were not vocabulary eliciting. Each of the *shared book reading* conditions was compared with a one-to-one reading condition that did not incorporate any questioning. The same three books were read by the author to children in all three conditions in four sessions over 6 weeks. The study took place in Head Start classrooms in an urban county in northern Texas. Children, aged 2 years, 7 months to 5 years, 4 months, were matched based on their baseline scores on the Peabody Picture Vocabulary Test-III (PPVT-III), and 15 children were randomly assigned to each condition. The analytic sample comparing the group that experienced questioning practices that were vocabulary-eliciting to the no-questioning condition included 30 children, and the analytic sample comparing the group that experienced questioning practices that were not vocabulary eliciting to the no-questioning condition included 29 children.

Summary of studies meeting WWC group design standards with reservations

Justice et al. (2010) conducted a study that included a total of 84 teachers from Head Start, prekindergarten, and independent programs in Ohio and Virginia that enroll academically at-risk, low-income children. As part of Project STAR (Sit Together And Read), teachers were randomly assigned to one of three conditions, each lasting 30 weeks: high-dose shared reading with print referencing (four reading sessions per week); low-dose shared reading with print referencing (two sessions per week); and a business-as-usual read-aloud comparison condition (four sessions per week). This study compared the high-dose shared reading intervention with the business-as-usual read-aloud comparison condition.[11] In both conditions, teachers used the same 30 books and followed the same reading schedule, reading one book per week, during four whole-class reading sessions. Teachers in the intervention condition used verbal and nonverbal print-referencing techniques, and teachers in the comparison group read as usual. A random sample of children with parental consent was selected to participate. The analytic sample included 58 classrooms: 30 intervention classrooms with 159 children, and 28 comparison classrooms with 129 children. This study was a cluster randomized controlled trial that may have included children who joined study classrooms subsequent to random assignment, and the authors did not indicate that they were answering a cluster-level research question. Because the study demonstrated the equivalence of the intervention and comparison groups in the analytic sample, the study *meets WWC group design standards with reservations*.

Pollard-Durodola et al. (2011) examined the effect of *shared book reading* on comprehension. The study was conducted in 18 Head Start and public preschool classrooms in eight schools, serving low-income children and

located in two ethnically diverse cities in south central Texas. Eighteen teachers were randomly assigned to either the Project WORLD (Words of Oral Reading and Language Development) intervention (11 teachers) or the business-as-usual read-aloud comparison (seven teachers). In the 12-week Project WORLD intervention, teachers conducted daily 20-minute shared reading sessions with groups of nine or ten children. Sessions focused on one storybook and one informational book each week, with teachers explaining target vocabulary, asking questions, and engaging children in discussions about the books and related science themes. In the comparison group, teachers read books to the whole class, using their usual reading style, which typically involved reading either storybooks or informational texts to the entire class for approximately 12 minutes, with the majority of time spent reading the text. The analytic sample consisted of 125 children who met eligibility criteria of being proficient in English and having well-below-average vocabulary knowledge (indicated by scoring below the 30th percentile on the PPVT-III), with 69 in intervention classrooms and 56 in comparison classrooms. This study was a cluster randomized controlled trial that included children who joined study classrooms subsequent to random assignment, and the authors did not indicate that they were answering a cluster-level research question. Because the study demonstrated the baseline equivalence of the intervention and comparison groups in the analytic sample, the study *meets WWC group design standards with reservations*.

Effectiveness Summary

The WWC review of *shared book reading* for the Early Childhood Education topic area includes outcomes in eight domains: comprehension, language development, alphabetics, general reading achievement, cognition, fluency, mathematics, and social-emotional development. The eight studies of *shared book reading* that meet WWC group design standards reported findings in four of the eight domains: (a) comprehension, (b) language development, (c) alphabetics, and (d) general reading achievement. One of the eight studies, which measured outcomes in the alphabetics and comprehension domains, does not contribute to effectiveness ratings because the information necessary to calculate effect sizes was not reported. The findings below present the authors' estimates and WWC-calculated estimates of the size and statistical significance of the effects of *shared book reading* on preschool children for seven of the eight studies. For a more detailed description of the rating of effectiveness and extent of evidence criteria, see the WWC Rating Criteria on p. 41.

Summary of effectiveness for the comprehension domain

Four studies that meet WWC group design standards—three without reservations and one with reservations—reported findings in the comprehension domain.

Bochna (2010) examined three researcher-developed outcomes in the comprehension domain: Topic–Prompted Recall, Main Idea–Prompted Recall, and Main Idea–Free Recall. The study author reported one positive and statistically significant effect on the Topic–Prompted Recall measure, which was confirmed by the WWC. No effects were negative and statistically significant; therefore, the average effect was statistically significant. The WWC characterizes these study findings as a statistically significant positive effect.

Lamb (1986) did not find statistically significant effects of *shared book reading* on the Peabody Picture Vocabulary Test–Revised (PPVT-R), compared to two comparison conditions (read-aloud without interaction comparison and interaction without read-aloud comparison). However, the WWC-calculated effect of read-aloud with structured language interaction compared to the interaction without read-aloud comparison was positive and substantively important (0.28). The effect of read-aloud with structured language interaction compared to read-aloud without interaction was not large enough to be considered substantively important (i.e., at least 0.25). The WWC characterizes these findings as an indeterminate effect.

Pollard-Durodola et al. (2011) examined effects on two outcomes in the comprehension domain: PPVT-III and Expressive One-Word Picture Vocabulary Test (EOWPVT). The study authors did not find statistically significant effects on either of the comprehension outcomes when comparing *shared book reading* focused on building vocabulary and background knowledge to teachers' usual whole-class book reading style. The WWC-calculated effects were not large enough to be considered substantively important. The WWC characterizes these findings as an indeterminate effect.

Walsh (2010) examined the effects of two *shared book reading* conditions on two researcher-developed comprehension outcomes: Seasonal Word Production Game and Seasonal Word Comprehension Game. For the *shared book reading* intervention condition (using the group that experienced questioning practices that were not vocabulary-eliciting), the study author did not find statistically significant effects on either outcome when compared to book reading without questions. The WWC-calculated effects were not large enough to be considered substantively important. For the second *shared book reading* intervention condition (using the experienced questioning practices that were vocabulary-eliciting), the author also did not find statistically significant effects on either outcome when compared to book reading without questions; however, each of the WWC-calculated effects were large enough to be considered substantively important—with a negative effect on the Seasonal Word Comprehension Game (–0.34) and a positive effect on the Seasonal Word Production Game (0.46). The average effect was not large enough to be considered substantively important. The WWC characterizes these study findings as an indeterminate effect.

Thus, for the comprehension domain, one study showed a statistically significant positive effect, and three studies showed indeterminate effects. This results in a rating of mixed effects, with a small extent of evidence.

Table 3.1 Rating of effectiveness and extent of evidence for the comprehension domain

Rating of effectiveness	Criteria met
Mixed effects *Evidence of inconsistent effects.*	In the four studies that reported findings, the estimated impact of the intervention on outcomes in the *comprehension* domain was mixed: one study showed a statistically significant positive effect, and three studies showed indeterminate effects.
Extent of evidence	**Criteria met**
Small	Four studies that included 233 children in at least 12 schools reported evidence of effectiveness in the *comprehension* domain.

Summary of effectiveness for the language development domain

Four studies that meet WWC group design standards—three without reservations and one with reservations—reported findings in the language development domain.

Justice et al. (2010) did not find a statistically significant effect of *shared book reading* on a composite measure of language ability, constructed from subtests of the Clinical Evaluation of Language Fundamentals–Preschool: 2 (CELF-P: 2), compared to teachers reading as they normally would. The WWC-calculated effect was not large enough to be considered substantively important. The WWC characterizes this study finding as an indeterminate effect.

Lamb (1986) did not find statistically significant effects of *shared book reading* on children's scores on the Record of Oral Language (Clay et al., 1983), compared to two comparison conditions (read-aloud without interaction and interaction without read-aloud). The WWC-calculated effect of read-aloud with structured interaction, compared to read-aloud without interaction, was negative and substantively important (–0.52), but the effect compared to the interaction without read-aloud condition was not large enough to be considered substantively important (–0.08). Because the mean of these two effects is negative and not statistically significant but is substantively important (–0.30), the WWC characterizes these study findings as a substantively important negative effect.

Mason et al. (1990) did not find a statistically significant effect of *shared book reading* on the Test of Early Language Development (TELD), compared to literacy lessons without *shared book reading*. The WWC-calculated effect was not large enough to be considered substantively important; therefore, the WWC characterizes this study finding as an indeterminate effect.

Mautte (1991) did not find a statistically significant effect of repeated storybook reading with adult interaction, compared to repeated storybook reading without adult interaction, on the Preschool Language Scale (PLS). However, the WWC-calculated effect was positive and large enough to be considered substantively important (0.36). The WWC characterizes these study findings as a substantively important positive effect.

Thus, for the language development domain, one study showed a substantively important positive effect, two studies showed an indeterminate effect, and one study showed a substantively important negative effect. This results in a rating of mixed effects, with a medium to large extent of evidence.

Table 3.2 Rating of effectiveness and extent of evidence for the language development domain

Rating of effectiveness	Criteria met
Mixed effects *Evidence of inconsistent effects.*	In the four studies that reported findings, the estimated impact of the intervention on outcomes in the *language development* domain was mixed: one study showed a substantively important positive effect; two studies showed indeterminate effects; and one study showed a substantively important negative effect.
Extent of evidence	**Criteria met**
Medium to large	Four studies that included 576 children in at least 77 classrooms in at least four schools reported evidence of effectiveness in the *language development* domain.

Summary of effectiveness for the alphabetics domain

Two studies that meet WWC group design standards—one without reservations and one with reservations—reported findings that contribute to the effectiveness rating in the alphabetics domain.

Justice et al. (2010) did not find a statistically significant effect of *shared book reading* on children's scores on a composite measure of print knowledge created from the Preschool Word and Print Awareness (PWPA) Test, the Phonological Awareness Literacy Screening for Preschool (PALS-PreK) Upper-Case Alphabet Recognition Subtest, and the PALS-PreK Name Writing Subtest. The WWC-calculated effect was not large enough to be considered substantively important; therefore, the WWC characterizes these study findings as an indeterminate effect.

Lamb (1986) did not find statistically significant effects of *shared book reading* on children's scores on the Concepts About Print Test (Clay, 1979), compared to two comparison conditions (read-aloud with no interaction comparison and interaction without read-aloud comparison). The WWC-calculated effect of read-aloud with structured interaction compared to the interaction without read-aloud comparison (0.43) was positive and substantively important. However, the effect of read-aloud with structured interaction compared to read-aloud without interaction was not large enough to be considered substantively important. The WWC characterizes these study findings as an indeterminate effect.

Thus, for the alphabetics domain, two studies showed indeterminate effects, with no studies showing a statistically significant or substantively important effect, either positive or negative. This results in a rating of no discernible effects, with a small extent of evidence.

Table 3.3 Rating of effectiveness and extent of evidence for the alphabetics domain

Rating of effectiveness	Criteria met
No discernible effects *No affirmative evidence of effects.*	In the two studies that reported findings, the estimated impact of the intervention on outcomes in the *alphabetics* domain was neither statistically significant nor large enough to be substantively important.
Extent of evidence	**Criteria met**
Small	Two studies that included 316 children in at least 60 classrooms reported evidence of effectiveness in the *alphabetics* domain.

Summary of effectiveness for the general reading achievement domain

One study that meets WWC group design standards without reservations reported findings in the general reading achievement domain.

Mautte (1991) did not find a statistically significant effect of repeated storybook reading with adult interaction, compared to repeated storybook reading without adult interaction, on the Test of Early Reading Ability (TERA), a measure of general reading achievement including both alphabetics and comprehension. The WWC-calculated effect size was not large enough to be considered substantively important according to WWC criteria. The WWC characterizes this study finding as an indeterminate effect.

Thus, for the general reading achievement domain, one study showed an indeterminate effect, with no statistically significant or substantively important effects, either positive or negative. This results in a rating of no discernible effects, with a small extent of evidence.

Table 3.4 Rating of effectiveness and extent of evidence for the general reading achievement domain

Rating of effectiveness	Criteria met
No discernible effects *No affirmative evidence of effects.*	In the one study that reported findings, the estimated impact of the intervention on outcomes in the *general reading achievement* domain was neither statistically significant nor large enough to be substantively important.
Extent of evidence	**Criteria met**
Small	One study that included 38 children in six classrooms in one school reported evidence of effectiveness in the *general reading achievement* domain.

References

Studies that meet WWC group design standards without reservations

Bochna, C. R. (2010). *The impact of instruction in text structure on listening comprehension in preschool age students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3378045)

Box, J. A., & Aldridge, J. (1993). Shared reading experiences and Head Start children's concepts about print and story structure. *Perceptual and Motor Skills, 77*(3), 929–930.

Additional source:

Box, J. A. (1991). *The effects of shared reading experiences on Head Start children's concepts about print and story structure* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9107738)

Lamb, H. A. (1986). *The effects of a read-aloud program with language interaction* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8616894)

Mason, J. M., Kerr, B. M., Sinha, S., & McCormick, C. (1990). Shared book reading in an Early Start program for at-risk children. *National Reading Conference Yearbook, 39,* 189–198.

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* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9115887)

Walsh, B. A. (2010). *Novel word learning of preschoolers enrolled in Head Start regular and bilingual classrooms: Impact of adult vocabulary noneliciting questions during shared storybook reading* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3384573)

Studies that meet WWC group design standards with reservations

Justice, L. M., McGinty, A. S., Piasta, S. B., Kaderavek, J. N., & Fan, X. (2010). Print-focused read-alouds in preschool classrooms: Intervention effectiveness and moderators of child outcomes. *Language, Speech, and Hearing Services in Schools, 41*(4), 504–520.

Additional sources:

Justice, L. M., Kaderavek, J. N., Fan, X., Sofka, A., & Hunt, A. (2009). Accelerating preschoolers' early literacy development through classroom-based teacher–child storybook reading and explicit print referencing. *Language, Speech, and Hearing Services in Schools, 40*(1), 67–85.

Piasta, S. B., Justice, L. M., McGinty, A. S., & Kaderavek, J. N. (2012). Increasing young children's contact with print during shared reading: Longitudinal effects on literacy achievement. *Child Development, 83*(3), 810–820.

Pollard-Durodola, S., Gonzalez, J. E., Simmons, D. C., Kwok, O., Taylor, A. B., Davis, M. J., … Simmons, L. (2011). The effects of an intensive shared book-reading intervention for preschool children at risk for vocabulary delay. *Exceptional Children, 77*(2), 161–183.

Studies that do not meet WWC group design standards

Blewitt, P., Rump, K. M., Shealy, S. E., & Cook, S. A. (2009). Shared book reading: When and how questions affect young children's word learning. *Journal of Educational Psychology, 101*(2), 294–304. The study does not meet WWC group design standards because equivalence of the analytic intervention and comparison groups was necessary and not demonstrated.

Irlen, S. M. (2003). *The impact of video viewing and retelling on preliterate children's narrative comprehension* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3088967) The study does not meet WWC group design standards because equivalence of the analytic intervention and comparison groups prior to the intervention was necessary and not demonstrated and the measures of effectiveness could not be attributed solely to the intervention.

Kertoy, M. K. (1994). Adult interactive strategies and the spontaneous comments of preschoolers during joint storybook readings. *Journal of Research in Childhood Education, 9*(1), 58–67. The study does not meet WWC group design standards because the eligible outcomes did not meet WWC requirements.

Morrow, L. M. (1988). Young children's responses to one-to-one story readings in school settings. *Reading Research Quarterly, 23*(1), 89–107. The study does not meet WWC group design standards because the eligible outcomes did not meet WWC requirements.

Murray, B. A., Stahl, S. A., & Ivey, M. G. (1996). Developing phoneme awareness through alphabet books. *Reading & Writing, 8*(4), 307–322. The study does not meet WWC group design standards because the measures of effectiveness could not be attributed solely to the intervention.

Additional source:

Murray, B. A., Stahl, S. A., & Ivey, M. G. (1993, December). *Developing phonological awareness through alphabet books.* Paper presented at the annual meeting of the National Reading Conference, Charleston, SC.

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

Barnes, E. M. (2013). *Head Start preschool teachers' commenting practices during shared book reading sessions: Describing learning opportunities for children with varying vocabulary abilities* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3578601) The study is ineligible for review because it did not use an eligible design.

Beauchat, K. A. (2008). *Making the most of shared reading in preschool* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3325480) The study is ineligible for review because it was out of the scope of the protocol.

Butler, M. C. (2013). *Implementation of evidence-based book-reading strategies by Head Start teachers: Benefits of professional development and effect on children's literacy outcomes* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3512230) The study is ineligible for review because it was out of the scope of the protocol.

Chambers, B., Cheung, A., Slavin, R. E., Smith, D., & Laurenzano, M. (2010). *Effective early childhood education programs: A systematic review.* Baltimore, MD: Johns Hopkins University, Center for Research and Reform in Education. http://files.eric.ed.gov/fulltext/ED527643.pdf. The study is ineligible for review because it did not use an eligible design.

Cooper, A. R. (2011). *Effects of group size on 4K teachers' use of decontextualized language during shared book reading with groups of children* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3325480) The study is ineligible for review because it did not use an eligible design.

Diehl, S. F., & Vaughn, B. (2010). Clinical discourse and engagement during shared storybook reading in preschool groups. *Seminars in Speech and Language, 31*(2), 111–121. The study is ineligible for review because it was out of the scope of the protocol.

Early childhood education: Top-rated programs. (2012). Retrieved from http://www.bestevidence.org The study is ineligible for review because it did not use an eligible design.

Ezell, H. K., & Justice, L. M. (2000). Increasing the print focus of adult-child shared book reading through observational learning. *American Journal of Speech-Language Pathology, 9*(1), 36–47. The study is ineligible for review because it was out of the scope of the protocol.

Fischer, R. L. (2004). *A naturalistic study of urban children's curiosity, engagement and persistence during shared book reading with teenage mothers* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3128532) The study is ineligible for review because it was out of the scope of the protocol.

Gonzalez, J. E., Pollard-Durodola, S., Simmons, D. C., Taylor, A. B., Davis, M. J., Kim, M., & Simmons, L. (2011). Developing low-income preschoolers' social studies and science vocabulary knowledge through content-focused shared book reading. *Journal of Research on Educational Effectiveness, 4*(1), 25–52. doi:10.1080/19345747.2010.487927 The study is ineligible for review because it was out of the scope of the protocol.

Greenberg, S. (2007). *The reading behaviors of preschool teachers during shared book reading with a narrative vs. expository text* (Unpublished doctoral dissertation). Pennsylvania State University, University Park. The study is ineligible for review because it was out of the scope of the protocol.

Hermanns, C. B. (2011). *Leveling the playing field: Investigating vocabulary development in Latino preschool-age English language learners* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3446324) The study is ineligible for review because it did not use a sample aligned with the protocol.

Jones, C. B. (1987). *The effect of oral reading by senior citizens on the oral language and readiness skills of language-delayed prekindergarten children* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8713620) The study is ineligible for review because it did not use an eligible design.

Korat, O., Shamir, A., & Heibal, S. (2013). Expanding the boundaries of *shared book reading*: E-books and printed books in parent–child reading as support for children's language. *First Language, 33*(5), 504–523. doi:10.1177/0142723713503148 The study is ineligible for review because it was out of the scope of the protocol.

LaCour, M. M. (2010). *The impact of a caregiver workshop regarding storybook reading on prekindergarten children's readiness for reading* (Unpublished doctoral dissertation). Liberty University, Lynchburg, VA. The study is ineligible for review because it was out of the scope of the protocol.

Lafferty, A. E. (2006). *Promoting emergent literacy skills with Spanish-speaking children: Examining children's outcomes using a Spanish-English book reading intervention* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3210159) The study is ineligible for review because it did not use a sample aligned with the protocol.

Liboiron, N., & Soto, G. (2006). Shared storybook reading with a student who uses alternative and augmentative communication: A description of scaffolding practices. *Child Language Teaching and Therapy, 22*(1), 69–95. doi:10.1191/0265659006ct298oa The study is ineligible for review because it did not use an eligible design.

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. The study is ineligible for review because it was out of the scope of the protocol.

Makin, L., & Whiteman, P. (2006). Young children as active participants in the investigation of teaching and learning. *European Early Childhood Education Research Journal, 14*(1), 33–41. doi:10.1080/13502930685209791 The study is ineligible for review because it did not use an eligible design.

Martin, N. M., Duke, N. K., Christ, T., & Reynold, J. M. (2010). Research worth knowing about. *Michigan Reading Journal, 42*(3), 48–52. The study is ineligible for review because it did not use an eligible design.

McCormick, C. E., & Mason, J. M. (1989). Fostering reading for Head Start children with Little Books. In J. Allen & J. M. Mason (Eds.), *Risk makers, risk takers, risk breakers: Reducing the risks for young literacy learners* (pp. 54–177). Portsmouth, NH: Heinemann. The study is ineligible for review because it was out of the scope of the protocol.

Additional resource:

McCormick, C. E., & Mason, J. M. (1986). *Use of Little Books at home: A minimal intervention strategy that fosters early reading.* Center for the Study of Reading: Technical Report, No. 338 (ED 314742). http://files.eric.ed.gov/fulltext/ED314742.pdf.

Milburn, T. F. (2011). *The efficacy of a professional development program to enhance preschool educators' ability to facilitate conversation during shared book reading* (Unpublished master's thesis). University of Ontario, Canada. The study is ineligible for review because it was out of the scope of the protocol.

Milburn, T. F., Girolametto, L. E., Weitzman, E., & Greenberg, J. (2014). Enhancing preschool educators' ability to facilitate conversations during shared book reading. *Journal of Early Childhood Literacy, 14*(1), 105–140. doi:10.1177/1468798413478261 The study is ineligible for review because it was out of the scope of the protocol.

Morgan, A. (2005). Shared reading interactions between mothers and pre-school children: Case studies of three dyads from a disadvantaged community. *Journal of Early Childhood Literacy, 5*(3), 279–304. doi:10.1177/1468798405058689 The study is ineligible for review because it was out of the scope of the protocol.

National Institute for Literacy. (2008). *Developing early literacy: Report of the National Early Literacy Panel. A scientific synthesis of early literacy development and implications for intervention*. Jessup, MD: Author. The study is ineligible for review because it did not use an eligible design.

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), 495–513. The study is ineligible for review because it did not use an eligible design.

Piasta, S. B., Dynia, J. M., Justice, L. M., Pentimonti, J. M., Kaderavek, J. N., & Schatschneider, C. (2010). Impact of professional development on preschool teachers' print references during shared read alouds: A latent growth curve analysis. *Journal of Research on Educational Effectiveness, 3*(4), 343–380. doi:10.1080/19345747.2010.482177 The study is ineligible for review because it was out of the scope of the protocol.

Pile, E. J. S., Girolametto, L., Johnson, C. J., Chen, X., & Cleave, P. L. (2010). Shared book reading intervention for children with language impairment: Using parents-as-aides in language intervention. *Canadian Journal of Speech-Language Pathology and Audiology, 34*(2), 96–109. The study is ineligible for review because it did not use a sample aligned with the protocol.

Pillinger, C., & Wood, C. (2013). A small-scale comparison of the relative impact of dialogic and shared book reading with an adult male on boys' literacy skills. *Journal of Early Childhood Literacy, 13*(4), 555–572. doi:10.1177/1468798413491975 The study is ineligible for review because it was out of the scope of the protocol.

Pillinger, C., & Wood, C. (2014). Pilot study evaluating the impact of dialogic reading and shared reading at transition to primary school: Early literacy skills and parental attitudes. *Literacy.* doi:10.1111/lit.12018 The study is ineligible for review because it was out of the scope of the protocol.

Pollard-Durodola, S. D., Mathes, P. G., Vaughn, S., Cardenas-Hagan, E., & Linan-Thompson, S. (2006). The role of oracy in developing comprehension in Spanish-speaking English language learners. *Topics in Language Disorders, 26*(4), 365–384. The study is ineligible for review because it did not use a sample aligned with the protocol.

Ratliff, J. L. (1987). *Explicit instruction in story structure: Effects on preschoolers' listening comprehension* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8625352) The study is ineligible for review because it was out of the scope of the protocol.

Reese, E., & Cox, A. (1999). Quality of adult book reading affects children's emergent literacy. *Developmental Psychology, 35*(1), 20–28. The study is ineligible for review because it was out of the scope of the protocol.

Reynolds, K. P. (2011). *Approaches to difficult words encountered during shared book reading: The impact of child shyness and reading ability* (Unpublished master's thesis). University of Guelph, Ontario. The study is ineligible for review because it did not use an eligible design.

Schickedanz, J. A., & McGee, L. M. (2010). The NELP report on shared story reading interventions (Chapter 4): Extending the story. *Educational Researcher, 39*(4), 323–329. The study is ineligible for review because it did not use an eligible design.

Shanahan, T., & Lonigan, C. J. (2010). The National Early Literacy Panel: A summary of the process and the report. *Educational Researcher, 39*(4), 279–285. The study is ineligible for review because it did not use an eligible design.

Shulstad, V. (2008). *The relationship between reports of shared book reading in the home and outcomes on the individual growth and development indicators among preschool children* (Doctoral dissertation, Minnesota State University, Moorhead). The study is ineligible for review because it was out of the scope of the protocol.

Silverman, R., Crandell, J. D., & Carlis, L. (2013). Read alouds and beyond: The effects of read aloud extension activities on vocabulary in Head Start classrooms. *Early Education & Development, 24*(2), 98–122. doi:10.1080/10409289.2011.649679 The study is ineligible for review because it was out of the scope of the protocol.

Snyder, C. K. C. (2013). *Effects of training on early childhood special education paraeducators' use of early literacy strategies during book reading* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3553235) The study is ineligible for review because it did not use a sample aligned with the protocol.

Stephenson, K. A. (2011). *Environmental, motivational, and cognitive predictors of emergent literacy and reading skills* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. NR70851) The study is ineligible for review because it did not use an eligible design.

Yeager Pelatti, C., Pentimonti, J. M., & Justice, L. M. (2014). Methodological review of the quality of reach out and read: Does it "work"? *Clinical Pediatrics, 53*(4), 343–350. doi:10.1177/0009922813507995 The study is ineligible for review because it did not use an eligible design.

Zucker, T. A., Cabell, S. Q., Justice, L. M., Pentimonti, J. M., & Kaderavek, J. N. (2013). The role of frequent, interactive prekindergarten shared reading in the longitudinal development of language and literacy skills. *Developmental Psychology, 49*(8), 1425–1439. The study is ineligible for review because it did not use an eligible design.

Appendix A.1: Research details for Bochna, 2010

Bochna, C. R. (2010). *The impact of instruction in text structure on listening comprehension in preschool age students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3378045)

Table A1. Summary of findings

Meets WWC group design standards without reservations

		Study findings	
Outcome domain	Sample size	Average improvement index *(percentile points)*	Statistically significant
Comprehension	36 preschool children	+18	Yes

Setting

The study was conducted in three Head Start centers in central Pennsylvania.

Study sample

Children ranged in age from 3–5 years at pretest. The children all qualified for full-day, full-year Head Start services based on their socioeconomic status. Forty-five children attending Head Start at three locations were randomly assigned to either the intervention condition (21 children) or the comparison condition (24 children). Within each condition, small groups were formed to participate in either the *shared book reading* condition or the comparison condition; there were five groups in each condition. The analytic sample includes 36 children—17 in the intervention condition and 19 in the comparison condition.

Intervention group

Children in the intervention group participated in 19 *shared book reading* sessions, each of which lasted approximately 10–12 minutes, with either the study author or one of two trained graduate students. The *shared book reading* sessions, which involved explicit instruction and questioning, focused on teaching children to identify the topic and main idea in expository texts.

The first 10 *shared book reading* sessions focused on identifying the topic. The adult read one book during each session. A total of five books were read, with each book being read twice. Sessions began with discussion focused on defining what a topic is. During reading, the adult paused periodically to ask children to identify the topic of the current book. After reading, the adult reviewed the definition of topic and asked children to identify the topic of the book.

The next nine *shared book reading* sessions focused on identifying the main idea. One book was read during each session. A total of five different books were read, with four of the five books being read twice. Before reading, there was a discussion of the concept of main idea and a review of what a topic is. During reading, the adult stopped periodically to ask children about the most important information in what was read. After reading, children were asked to identify the topic and main idea in the book.

Comparison group

Children in the comparison group participated in 19 sessions, in which the same adult read the same books as those in the intervention condition; however, there was no explicit instruction in identifying the topic or the main idea. Each book was read from start to finish without pause for discussion. The focus was on children listening to the books instead of talking about them.

Outcomes and measurement

In the comprehension domain, three researcher-developed outcome measures were used to assess children's understanding of the topic and main idea in expository text. The three outcomes—Main Idea–Free Recall; Topic–Prompted Recall; and Main Idea–Prompted Recall—were collected as part of a individually-administered book read-aloud activity. The assessment involved the author or one of two trained graduate students reading the child an unfamiliar book straight through without discussion. Following the read-aloud, the adult asked the child to tell everything he or she remembered about the story, scoring the number of idea units recalled for the Main Idea–Free Recall measure. Following the free recall, the adult prompted the child to tell the topic of the book for the Topic–Prompted Recall measure, and prompted the child to tell "what the book mostly tells about [topic]" for the Main Idea–Prompted Recall measure. The assessments were administered prior to the start of the intervention (pretest) and immediately following the 19 reading sessions (posttest). For a more detailed description of these outcome measures, see Appendix B.

Support for implementation

No training was provided for implementing the intervention.

Appendix A.2: Research details for Box & Aldridge, 1993

Box, J. A., & Aldridge, J. (1993). Shared reading experiences and Head Start children's concepts about print and story structure. *Perceptual and Motor Skills, 77*(3), 929–930.

Additional source:

Box, J. A. (1991). *The effects of shared reading experiences on Head Start children's concepts about print and story structure* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9107738)

Table A2. Summary of findings

Meets WWC group design standards without reservations

Outcome domain	Sample size	Study findings	
		Average improvement index *(percentile points)*	Statistically significant
Comprehension	72 preschool children	na	na
Alphabetics	72 preschool children	na	na[12]

na = not applicable.

Setting

The study took place in 15 classrooms located in Head Start centers and public schools in one geographic area.

Study sample

A total of 15 classrooms were chosen by the Head Start director from a total of 33 classrooms. Within each classroom, five children were randomly selected to participate in the study, and each class group was randomly assigned to one of three conditions—*shared book reading* (intervention), business-as-usual general classroom instruction (comparison), and movement activities (placebo). Classrooms had 18–22 children each, aged 3–5 years old. The 75 children selected for the study ranged in age from 4 years, 1 month to 4 years, 11 months. The children in all three conditions were similar with regard to socioeconomic status, as all were eligible for the Head Start program. All teachers involved in the study had the same level of education; high school graduates who had completed the Child Development Associate (CDA) National Credentialing Program requirements, without attaining a bachelor's degree in education.

Intervention group

The *shared book reading* intervention involved the classroom teacher reading with a group of five children for 10–15 minutes a day, 4 days per week, for 8 weeks. The *shared book reading* involved three phases: discovery, exploration, and independent experience and expression. The discovery phase took place for 2 days and involved the teacher introducing new books and encouraging children to chime in on repetitive sections, fill in missing words, and suggest possible story outcomes. The exploration phase took place during the rest of the week, beginning on the second day, and involved the teacher re-reading familiar books. Unison participation was common in this phase, which focused on teaching children story structure and relevant reading strategies. The third phase involved independent opportunities for children to read familiar books with the teacher outside of the small group reading experience.

Comparison group

There were two comparison conditions: (a) business-as-usual general classroom instruction, during which children had access to social studies and science units, as well as their usual learning centers, such as art, library, housekeeping, math, and language arts activities; and (b) movement activities, which did not involve literacy instruction. The children took part in these activities for 10–15 minutes a day, 4 days per week, for 8 weeks. The business-as-usual practice may have included some read-aloud activities as part of typical instruction, but structured interaction focused on the text was not explicitly used. The teacher led children in movement activities that they had not engaged in previously during the school year.

Outcomes and measurement

In the alphabetics domain, the Early School Inventory-Preliteracy/Part A Print Concepts was used to assess print awareness. Children were shown cards with pictures and/or print and asked to point to pictures or select words to demonstrate concepts of print.[13] In the comprehension domain, the Early School Inventory-Preliteracy/Part B Story Structure was administered, which measures children's ability to retell a familiar story by including specific elements necessary for the story to be complete. For a more detailed description of these outcome measures, see Appendix B.

Support for implementation

The author conducted centrally-located group training and demonstration sessions with the teachers in the shared reading intervention group and with the teachers in the movement activities placebo group. For intervention teachers, charts were provided with guidelines to follow during shared reading with children. Teachers in the placebo group were instructed to follow the guidelines on each of the movement records. Teachers in the intervention and placebo groups were monitored by the author five times during the 8-week intervention period to ensure that they adhered to the guidelines. Teachers in the comparison group did not require special instruction or support for implementation, because children in this condition were receiving the usual classroom instruction.

Appendix A.3: Research details for Lamb, 1986

Lamb, H. A. (1986). *The effects of a read-aloud program with language interaction* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8616894)

Table A3. Summary of findings

Meets WWC group design standards without reservations

Outcome domain	Sample size	Study findings	
		Average improvement index *(percentile points)*	Statistically significant
Comprehension	28 preschool children	+6	No
Language development	28 preschool children	–12	No
Alphabetics	28 preschool children	+4	No

Setting

The study took place in one day care center in Tallahassee, Florida.

Study sample

The day care center served primarily minority and low socioeconomic status (SES) families. The children in the study were aged 3–5 years, from a minority racial group, and mostly from low SES families. The 36 children were randomly assigned to either the intervention group—read-aloud with structured language interaction (10 children)—or one of three comparison groups: read-aloud with no interaction (9 children); interaction with no read-aloud (9 children); and no contact (8 children).

Intervention group

The read-aloud with structured language interaction condition involved the study author reading a preselected book with a small group of children, using verbal and nonverbal interactions before, during, and after the reading that focused on the book, story, and children's related experiences. The author met with the intervention group on a daily basis for 10 weeks to deliver the intervention.

Comparison group

There were three comparison conditions. The read-aloud with no interaction condition involved the author reading to a small group of children using the same books as in the intervention condition with no interaction initiated by the author. The interaction-only condition involved no read-aloud activities; the author engaged a small group of children in art activities, providing comments and discussing the activities. The author met with each of these comparison groups on a daily basis for 10 weeks. The third comparison group condition was business-as-usual classroom instruction, in which children had no contact with the author. The business-as-usual general classroom instruction condition was excluded from this intervention report because contrasts against this group do not meet WWC standards.[14]

Outcomes and measurement

In the comprehension domain, the author used the PPVT-R, which measures receptive vocabulary. In the alphabetics domain, the author used the Concepts About Print Test (Clay, 1979), which measures knowledge about book orientation, print convention, concepts of words and punctuation, and relationship between print and meaning. In the language development domain, the author used the Record of Oral Language (Clay et al., 1983), in which children repeat carefully constructed sentences. For a more detailed description of these outcome measures, see Appendix B.

Support for implementation

The author implemented the intervention and two of the comparison conditions. No support or training was provided.

Appendix A.4: Research details for Mason et al., 1990

Mason, J. M., Kerr, B. M., Sinha, S., & McCormick, C. (1990). Shared book reading in an Early Start program for at-risk children. *National Reading Conference Yearbook, 39,* 189–198.

Table A4. Summary of findings

Meets WWC group design standards without reservations

Outcome domain	Sample size	Study findings	
		Average improvement index *(percentile points)*	Statistically significant
Language development	232 four-year-olds	+3	No

Setting

The study was conducted in 12 classes within two Early Start schools in a mid-sized urban setting. Early Start is a half-day developmental program in Illinois for 4-year-old children at risk for school failure, identified based on developmental screening, home visits, and family interviews.

Study sample

In the two schools, there were a total of six classrooms in the study, each with a morning and an afternoon class. In five of the classrooms, the same lead teacher taught both the morning and afternoon classes. In one classroom, a lead teacher taught the morning class and an aide taught the afternoon class. Therefore, a total of seven teachers participated in the study. Within each classroom, classes were randomized: one to the intervention condition and one to the comparison condition, so five teachers taught classes in both conditions. There were six intervention classes (three morning and three afternoon) and six comparison classes (three morning and three afternoon). At baseline, there were 240 four-year-old children: 123 intervention and 117 comparison. The analytic sample included 232 four-year-old children: 115 intervention and 117 comparison. In the intervention group, 45% of the children were female, compared to 49% in the comparison group. The study authors indicated that the groups had approximately equal numbers of African-American and White children, and most were from families with low socioeconomic status (SES).

Intervention group

Each classroom teacher used Little Books as a small group activity from mid-September to the end of the school year. For 28 weeks, the intervention was implemented as follows. Each week, the teacher focused on a different Little Book. On Monday, Wednesday, and Friday of a given week, the teacher read the book aloud with a small group of children and encouraged discussion of the book; children were also encouraged to read the book with the teacher, to each other, and by themselves as they became comfortable. The teacher gradually corrected mistakes. Friday culminated with a related activity, and children were each given a copy of the book to take home and share with their families.

Comparison group

The comparison group did not have access to Little Books; they received instruction in all other literacy lessons and activities given to the intervention group. However, the common literacy lessons were slightly longer for the comparison group, as teachers in the intervention condition decreased the time spent on the standard lessons and activities to make time for Little Books.

Outcomes and measurement

In the language development domain, the author used the TELD, which was administered in September/October (baseline) and again in April (posttest) of the school year.[15] For a more detailed description of this outcome measure, see Appendix B.

Support for implementation

Two workshops and a follow-up visit were offered to teachers. In May of the prior school year, teachers and school personnel were introduced to Little Books and procedures for their use. In September, a workshop was held and follow-up visits were made when teachers began using Little Books to ensure fidelity.

Appendix A.5: Research details for Mautte, 1991

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* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9115887)

Table A5. Summary of findings

Meets WWC group design standards without reservations

		Study findings	
Outcome domain	Sample size	Average improvement index *(percentile points)*	Statistically significant
Language development	38 four-year-olds	+18	No
General reading achievement	38 four-year-olds	–8	No

Setting

The study was conducted in one early childhood education center in the Hillsborough County School System in the inner-city section of Tampa, Florida.

Study sample

The study focused on all 4-year-old children in the prekindergarten classes in the At-Risk Program, which serves children based on educational and economic need (i.e., families receiving federal assistance) or identified as suffering from abuse and neglect. The center had six prekindergarten classes with a total of 148 children, aged 2–4 years old. All 66 children who were 4 years old were included in the sample at baseline. The sample was grouped according to developmental level. Those above the median score on the district's Children's Inventory of Learning Development (CHILD) Test were considered "Average Development Level," and those below the median were considered "Delayed Development Level." Within each level, children were randomly assigned to either repeated storybook reading *with* adult interaction (intervention), repeated storybook reading *without* adult interaction (Comparison 1), or business-as-usual prekindergarten classroom instruction (Comparison 2).[16] At random assignment, there were 22 children in each of the three conditions. At posttest, there were 18 children in the intervention group, 20 children in the Comparison 1 group, and 15 children in the Comparison 2 group. This sample was comprised of 27 girls and 26 boys; 87% of the children were African American.

Intervention group

The *shared book reading* intervention, referred to as "adult-interaction repeated storybook reading" by the study author, was implemented as a pull-out program. The intervention involved the author reading Big Book storybooks aloud to children, one book each week, for a total of 20 weeks. The same Big Book was read three times during the week. Interaction occurred before, during, and after the book reading. Before reading, the author introduced the book, encouraged children to generate predictions about the story based on title and pictures, asked questions that related the children' lives and the story, and established a purpose for listening to the story. During the reading, the author showed illustrations, indicated the correspondence between spoken and written words with a pointer, clarified or explained the text, asked inferential questions or questions related to the purpose for listening, and encouraged children to generate predictions about the story content. After reading, the instructor encouraged discussion, asked questions to generate evaluative responses, and related the story concepts to children's lives. On subsequent readings, the children were asked to recall key aspects of the story and to participate in the reading (chanting refrains or filling in predictable phrases). Each session lasted for a maximum of 25 minutes and was conducted in the morning during the prekindergarten classrooms' large block of activity time.

Comparison group

The comparison condition was repeated storybook reading without adult interaction, which was a pull-out program in which the author read the same 20 Big Book storybooks used in the intervention condition. Big Book storybooks were read as written, without adult interaction. Books were read three times per week, one book per week, for 20 weeks. Each session lasted a maximum of 25 minutes and was conducted in the morning during the prekindergarten classrooms' large block of activity time. The author did not comment, ask questions, or answer questions before or during book reading. Although the author did not initiate any interaction with children, the author would answer children's questions after the reading was completed.

Outcomes and measurement

In the language development domain, the author used the PLS, a standardized measure of language development. In the general reading achievement domain, the author used the TERA, a standardized test of prereading, including environmental contexts, vocabulary, listening comprehension, alphabetics, and print awareness. Posttest assessments were administered at the end of the 20-week intervention period. For a more detailed description of these outcome measures, see Appendix B.

Support for implementation

The intervention and comparison storybook reading conditions were both implemented by the author, who audio-recorded each reading session and reviewed the tapes to monitor consistency of reading style in all small groups within each condition. In addition, reading instructors at the University of South Florida reviewed the tapes to assess fidelity to procedures and scripts.

Appendix A.6: Research details for Walsh, 2010

Walsh, B. A. (2010). *Novel word learning of preschoolers enrolled in Head Start regular and bilingual classrooms: Impact of adult vocabulary noneliciting questions during shared storybook reading* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3384573)

Table A6. Summary of findings

Meets WWC group design standards without reservations

Outcome domain	Sample size	Study findings	
		Average improvement index *(percentile points)*	Statistically significant
Comprehension	30 preschool children	+4	No

Setting

The study was conducted in Head Start classrooms in an urban county in northern Texas.

Study sample

The sample included 45 children in Head Start classrooms (20 female, 25 male), aged 3–5 years old, with a mean age of 4 years, 3 months. The children were ranked and grouped into triads based on their scores on the PPVT-III. Within each triad, children were randomly assigned to one of three conditions: the use of questioning practices that were vocabulary-eliciting, the use of questioning practices that were not vocabulary eliciting, and no questions (comparison). As a result, each condition had an assigned sample of 15 children. The analytic sample included 44 children: 15 in the vocabulary eliciting condition, 14 in the vocabulary non-eliciting condition, and 15 in the comparison condition.

Intervention group

The two intervention conditions—the use of questioning practices that were vocabulary-eliciting and the use of questioning practices that were not vocabulary-eliciting—included four sessions conducted over a 6-week period. In the first three sessions, children met with the adult reader for one-to-one *shared book reading* and were asked six questions during the reading related to the story. Three age-appropriate storybooks were designed for the experiment. Two of the three books were read in each session. Nine words, each appearing twice, were embedded in the stories. In the fourth session, all three stories were read. In the condition that used vocabulary-eliciting questions, children were asked questions that did not contain the target words but required target words as answers (e.g., "what are these [point to skis]?"). In the condition that used questioning practices that were not vocabulary-eliciting, children were asked questions that contained the target words and did not require use of the target words in the answers (e.g., "where are the skis?").

Comparison group

The comparison condition involved an adult reading one-to-one with a child, using the same three storybooks as in the two intervention conditions. As in the intervention conditions, the books were read to children at four sessions over a 6-week period. The adult read the book to the child without asking any questions.

Outcomes and measurement

In the comprehension domain, two researcher-designed vocabulary measures were used: Seasonal Word Production Game, in which children verbally provided labels for pictures of the nine target words, with one point awarded for each correct response; and Seasonal Word Comprehension Game, in which children heard each of the nine target words and selected the corresponding pictures from among four choices, with one point earned for each correct answer. For a more detailed description of these outcome measures, see Appendix B. One-week follow-up results are reported in Appendix D.

Support for implementation

The intervention and data collection were carried out by the author and one research assistant (a doctoral candidate). During a 1-hour training session, the author trained the assistant to implement the intervention.

Appendix A.7: Research details for Justice et al., 2010

Justice, L. M., McGinty, A. S., Piasta, S. B., Kaderavek, J. N., & Fan, X. (2010). Print-focused read-alouds in preschool classrooms: Intervention effectiveness and moderators of child outcomes. *Language, Speech, and Hearing Services in Schools, 41*(4), 504–520.

Additional source[17]:

Piasta, S. B., Justice, L. M., McGinty, A. S., & Kaderavek, J. N. (2012). Increasing young children's contact with print during shared reading: Longitudinal effects on literacy achievement. *Child Development, 83*(3), 810–820.

Table A7. Summary of findings **Meets WWC group design standards with reservations**

Outcome domain	Sample size	Study findings	
		Average improvement index *(percentile points)*	Statistically significant
Language development	278 preschool children	+3	No
Comprehension	288 preschool children	+9	No

Setting

The study was conducted in 59 preschool classrooms in Ohio and Virginia.

Study sample

Preschool classrooms were drawn from a variety of settings, all of which prioritized academically at-risk children for enrollment. The classrooms included: 23 affiliated with Head Start, 19 subsidized prekindergarten programs, 12 independent programs that accepted vouchers, and five early childhood special education programs. Lead preschool teachers were recruited to participate in the study at information sessions presented at early childhood organizations. Two waves of teachers (84 total) were randomly assigned as part of Project STAR to one of three conditions: a high-dose shared reading with print referencing condition, a low-dose shared reading with print referencing condition, and a comparison condition. Justice et al. (2010) compared the high-dose shared reading with print referencing (intervention) condition (31 classrooms) to the comparison condition (28 classrooms), for a total sample size of 59 classrooms at baseline. A random subsample of children for whom consent was obtained was included in the study, for a total baseline sample of 379 children (201 intervention and 178 comparison). The baseline sample of children was 54% female; 42% White, 37% African American, and 8% Hispanic. The

analytic sample included 58 classrooms—30 intervention and 28 comparison. The analytic sample included 278 children for the analysis of impacts on language development (151 intervention, 127 comparison) and 288 children for the analysis of impacts on alphabetics (159 intervention, 129 comparison).

Intervention group

The intervention condition, high-dose shared reading with print referencing, lasted 30 weeks with four whole-class reading sessions per week and no more than one session per day. Materials included a set of 30 books; a schedule for reading; and a description of the scope, sequence, and frequency of print-related targets to be addressed during each read-aloud. There were 15 defined print-knowledge objectives for each book. Each week, teachers would read the prescribed book, using verbal (e.g., questioning) and nonverbal (e.g., tracking print) references to address the print-knowledge targets for the book. After reading, each book was placed in the classroom library and not used for instruction or class reading during the study period. The low-dose shared reading with print referencing condition involved the same shared reading interactions and materials as the high-dose condition, but adults read with children twice per week instead of four times per week. No studies report immediate posttest results for the low-dose condition; however, 1-year and 2-year follow-up results are reported in Piasta et al. (2012) and are summarized in Appendix D.

Comparison group

Teachers in the comparison group also conducted a whole-class book-reading session four times weekly for 30 weeks. These teachers received the same set of 30 children's books and the same schedule for reading as those in the intervention group. They were instructed to simply read the books as they normally would. After reading, each book on the reading list was placed in the classroom library and not used for instruction or class reading during the study period.

Outcomes and measurement

To measure the alphabetics domain at posttest, a composite measure of print awareness was constructed from three standardized tests: the PWPA Test (Justice & Ezell, 2001; Justice, Bowles, & Skibbe, 2006), a structured task that examines children's print concepts, and two subtests of PALS–PreK (Invernizzi, Sullivan, Meier, & Swank, 2004)—the Upper-Case Alphabet Recognition Subtest, which asks children to name upper-case letters, and the Name Writing Subtest, which asks children to draw a self-portrait and then write their names on it. To measure the language development domain at posttest, the authors created a composite score based on three subtests of the CELF-P:2 (Wiig, Secord, & Semel, 2004): Sentence Structure, Word Structure, and Expressive Vocabulary. These subtests collectively measure language in the areas of vocabulary, syntax, and morphology and require approximately 15–20 minutes to administer. Assessments were administered to children in fall and spring of the school year.

At follow-up, 1 and 2 years after the end of the intervention, the authors used two standardized measures in the alphabetics domain: Woodcock-Johnson-III (WJ-III) Letter-Word Identification subtest and WJ-III Spelling subtest. At follow-up, the authors used two standardized measures in the comprehension domain: WJ-III Passage Comprehension subtest and Peabody Picture Vocabulary Test-4 (PPVT-4). For a more detailed description of these outcome measures, see Appendix B.

Support for implementation

Teachers in the shared reading with print-referencing intervention group received explicit directions and materials at the start of the academic year on how to implement a 30-week read-aloud program in their classrooms using a print-referencing style. Intervention teachers received 8 hours of professional development prior to the start of the school year, two feedback letters based on videos of their read-alouds, and another 3 hours of professional development mid-year.

Appendix A.8: Research details for Pollard-Durodola et al., 2011

Pollard-Durodola, S. D., Gonzalez, J. E., Simmons, D. C., Kwok, O., Taylor, A. B., Davis, M. J., … Simmons, L. (2011). The effects of an intensive shared book-reading intervention for preschool children at risk for vocabulary delay. *Exceptional Children, 77*(2), 161–183.

Table A8. Summary of findings **Meets WWC group design standards with reservations**

		Study findings	
Outcome domain	Sample size	Average improvement index *(percentile points)*	Statistically significant
Comprehension	125 preschool children	+2	No

Setting

The study was conducted in 18 classrooms from two school districts and one regional Head Start agency in two ethnically diverse cities in south central Texas. In one school district, the study took place in six classrooms in a half-day public preschool center. In the other school district, the study was conducted in nine full-day Head Start classrooms located in five different schools. In the regional Head Start agency, the study took place in three full-day Head Start classrooms located in two different schools. Parents did not have a choice of half-day or full-day setting, but were offered the program available in their district.

Study sample

Schools with a high percentage of children from low socioeconomic backgrounds were selected for the study. The 18 study teachers were randomly assigned to either the Project WORLD intervention condition (11 teachers) or the business-as-usual read-aloud comparison condition (7 teachers). There were 148 children with parental consent who met the eligibility criteria of having well-below-average vocabulary knowledge (indicated by scoring below the 30th percentile on the PPVT-III) and demonstrating English proficiency (as reported by their teacher); of these children, 81 were in intervention classrooms and 67 were in comparison classrooms. The analytic sample consisted of 125 children, with 69 in intervention classrooms and 56 in comparison classrooms. Children ranged in age from 4.0–5.3 years, with a mean age of 4.5 years. The sample of children was 53% female and 47% male, and all of them received free or reduced-price lunch. The racial/ethnic composition was 50% African American, 28% Latino, 14% Caucasian, and 8% Asian American.

Intervention group

The Project WORLD intervention was implemented for 20 minutes daily for 12 weeks. Teachers conducted the *shared book reading* sessions with groups of nine or ten children, while the remaining children were engaged in other activities supervised by the classroom paraprofessional. The lessons were organized around two science themes, Nature and Living Things, with each theme being taught for 6 weeks. Each week, the teaching unit focused on one theme-related topic, with two books per topic, one storybook and one informational book. Children were exposed to one book on the first 2 days and the second book on the next 2 days. The fifth day was a review day. Children received instruction on 68 target words taken from the 24 books in order to develop their background knowledge about the specific topics. During and after book reading, the teachers explained and reviewed target vocabulary, asked questions, and led discussions to engage children in the learning process.

Comparison group

Teachers in the comparison classrooms used their usual style when reading with the whole class, using books they selected from their classrooms or school library. Only three of their selections were the same books used in the intervention classrooms.

Outcomes and measurement

Outcomes included two measures in the comprehension domain: the PPVT-III, in which the children hear a word and select the correct answer from four pictures; and the EOWPVT, in which respondents are asked to name the objects or the actions illustrated in pictures.[18] For a more detailed description of these outcome measures, see Appendix B.

Support for implementation

Before the intervention, teachers in the intervention group received a half-day professional development session, in which the developers introduced Project WORLD and provided teachers with instructional materials. During the intervention, the teachers also met with the authors three times to report progress and discuss the obstacles they met with during implementation.

Appendix B: Outcome measures for each domain

Comprehension	
Early School Inventory–Preliteracy/ Part B: Story Structure Subtest	The Early School Inventory–Preliteracy/Part B: Story Structure Subtest measures children's abilities to retell a familiar story by including specific elements necessary for the story to be considered complete (as cited in Box & Aldridge, 1993).
Topic–Prompted Recall	Topic–Prompted Recall is a researcher-designed assessment that measures the ability to identify and recall the topic of a text in response to a question. In the assessment, after children are read a text, they are asked what the topic of the book was. The score of the measure is dichotomous (as cited in Bochna, 2010).
Main Idea–Prompted Recall	Main Idea–Prompted Recall is a researcher-designed assessment that measures the ability to identify and recall the main ideas of a text in response to a question. In the assessment, after children are read an expository text, they are asked what the book "mostly tells about." The score of the measure is the number of facts provided (as cited in Bochna, 2010).
Main Idea–Free Recall	Main Idea–Free Recall is a researcher-designed assessment that measures the ability to identify and recall details of a text with no prompting. In the assessment, after children are read an expository text, they are asked to tell the assessor everything that they remember about the text. The score of the measure is the number of facts provided (as cited in Bochna, 2010).
Peabody Picture Vocabulary Test–Revised (PPVT-R)	The PPVT-R (Dunn & Dunn, 1981) is a standardized measure of receptive vocabulary in which children demonstrate understanding of spoken words by pointing to pictures that represent their meanings (as cited in Lamb, 1986).
Peabody Picture Vocabulary Test–III (PPVT-III)	The PPVT-III (Dunn & Dunn, 1997) is a standardized measure of receptive vocabulary in which children demonstrate understanding of spoken words by pointing to pictures that represent their meanings (as cited in Pollard-Durodola et al., 2011).
Expressive One-Word Picture Vocabulary Test (EOWPVT)	The EOWPVT (Brownell, 2000) is a standardized measure of children's expressive vocabulary that asks them to produce the words describing objects and actions pointed to by an assessor (as cited in Pollard-Durodola et al., 2011).
Seasonal Word Comprehension Game	The Seasonal Word Comprehension Game is a researcher-designed measure of receptive vocabulary. Children hear nine target words and select the corresponding pictures from four choices. One point is awarded for each correct response (as cited in Walsh, 2010).
Seasonal Word Production Game	The Seasonal Word Production Game is a researcher-designed measure of expressive vocabulary. Children demonstrate their understanding of nine target words by naming them when an assessor points to their pictures. One point is awarded for each correct response (as cited in Walsh, 2010).
Woodcock-Johnson III (WJ-III) Tests of Achievement: Passage Comprehension subtest	The WJ-III is a nationally-normed, individually-administered battery of tests of achievement. The Passage Comprehension subtest measures reading comprehension and requires children to identify pictures related in meaning and to select a picture or produce a word that accurately completes a phrase or passage (as cited in Piasta et al, 2012).
Peabody Picture Vocabulary Test, Fourth Edition (PPVT-4)	The PPVT-4 (Dunn & Dunn, 2007) is a standardized measure of receptive vocabulary in which children demonstrate understanding of spoken words by pointing to pictures that represent their meanings (as cited in Piasta et al, 2012).
Language development	
Composite Measure of Language Ability	The composite measures children's language ability (including vocabulary, syntax, and morphology) by combining the scores from three subtests of the Clinical Evaluation of Language Fundamentals–Preschool: 2 (CELF-P:2) (Wiig, Secord, & Semel, 2004): (a) the Sentence Structure Subtest, which measures comprehension of complex sentence structures; (b) the Word Structure Subtest, which measures the ability to utilize morphological structures (e.g., verb tenses and pronouns); and (c) the Expressive Vocabulary Subtest, which measures the ability to produce the names of people, actions, and objects (as cited in Justice et al., 2010).
Record of Oral Language	The Record of Oral Language (Clay et al., 1983) is a non-standardized measure of children's oral language abilities that asks them to repeat sentences (as cited in Lamb, 1986).
Test of Early Language Development (TELD)	The TELD (Hresko, Reid, & Hammill, 1981) is a standardized measure of children's receptive and expressive language skills (as cited in Mason et al., 1990).
Preschool Language Scale (PLS)	The PLS (Zimmerman, Steiner, & Pond, 1979) is a standardized measure of children's receptive and expressive language (as cited in Mautte, 1991).

Alphabetics	
Composite Measure of Print Knowledge	The composite score measures children's print knowledge. The score is created by combining the scores from three tests: (a) The PWPA Test (Justice & Ezell, 2001; Justice, Bowles & Skibbe, 2006), which measures print knowledge by asking children during shared reading to indicate or explain 14 aspects of print (e.g., directionality of print); (b) the Phonological Awareness Literacy Screening for Preschool (PALS-PreK; Invernizzi, Sullivan, Meier, & Swank, 2004) Upper-Case Alphabet Recognition Subtest, which measures the ability to name upper-case letters presented in random order; and (c) the PALS-PreK Name Writing Subtest, which measures children's name-writing ability by having children write their names and rating the result for its correctness (as cited in Justice et al., 2010).
WJ-III Tests of Achievement: Letter-Word Identification subtest	The WJ-III is a nationally-normed, individually-administered battery of tests of achievement. The Letter-Word Identification subtest measures basic word reading skills and requires children to identify letters and read aloud individual words of increasing difficulty (as cited in Piasta et al, 2012).
WJ-III Tests of Achievement: Spelling subtest	The WJ-III is a nationally-normed, individually-administered battery of tests of achievement. The Spelling subtest measures early spelling skill and requires children to write letters and spell words of increasing difficulty (as cited in Piasta et al, 2012).
Concepts About Print Test	The Concepts About Print Test (Clay, 1979) measures knowledge of print (e.g., book orientation, concepts of words, punctuation, and the relationship between print and oral language) (as cited in Lamb, 1986).
Early School Inventory–Preliteracy/ Part A: Print Concepts Subtest	The Early School Inventory–Preliteracy/Part A: Print Concepts Subtest measures print awareness. Children are shown cards with pictures and/or print, and asked to point to the pictures or select words to demonstrate concepts of print (as cited in Box & Aldridge, 1993).
General reading achievement	
Test of Early Reading Ability (TERA)	The TERA is a standardized measure of early reading skills, including awareness and concepts, vocabulary, listening, comprehension, and alphabet knowledge (as cited in Mautte, 1991).

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

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			*p*-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
Bochna, 2010[a]								
Topic–Prompted Recall (dichotomous)	Preschool children	36	0.69 (na)	0.32 (na)	0.37	0.91	+32	< .01
Main Idea–Prompted Recall	Preschool children	36	0.84 (0.87)	0.37 (0.50)	0.47	0.66	+24	.43
Main Idea–Free Recall	Preschool children	36	0.91 (1.82)	0.95 (1.31)	–0.04	–0.03	–1	.21
Domain average for comprehension (Bochna, 2010)						**0.51**	**+19**	**Statistically significant**
Box & Aldridge, 1993[b]								
Early School Inventory–Preliteracy/Part B: Story Structure Subtest	Shared reading vs. usual instruction	48	4.37 (nr)	3.26 (nr)	1.11	nr	nr	nr
Early School Inventory–Preliteracy/Part B: Story Structure Subtest	Shared reading vs. movement activities	49	4.37 (nr)	4.01 (nr)	0.36	nr	nr	nr
Domain average for comprehension (Box & Aldridge, 1993)						**nr**	**nr**	**nr**
Lamb, 1986[c]								
Peabody Picture Vocabulary Test–Revised (PPVT-R)	Reading & interaction vs. reading without interaction	19	27.30 (14.03)	27.44 (18.23)	–0.14	–0.01	0	>.05
PPVT-R	Reading & interaction vs. interaction without reading	19	27.30 (14.03)	23.88 (8.51)	3.42	0.28	+11	> .05
Domain average for comprehension (Lamb, 1986)						**0.13**	**+5**	**Not statistically significant**
Domain average for comprehension (reading & interaction vs. reading without interaction; Lamb, 1986)						**–0.01**	**0**	**Not statistically significant**
Domain average for comprehension (reading & interaction vs. interaction without reading; Lamb, 1986)						**0.28**	**+11**	**Not statistically significant**
Pollard-Durodola et al., 2011[d]								
Peabody Picture Vocabulary Test–III (PPVT-III)	Preschool children	125	84.82 (12.51)	84.30 (14.00)	0.52	0.04	+2	**.80**
Expressive One-Word Picture Vocabulary Test (EOWPVT)	Preschool children	125	80.98 (8.98)	80.34 (10.93)	0.64	0.06	+3	**.70**
Domain average for comprehension (Pollard-Durodola et al., 2011)						**0.05**	**+2**	**Not statistically significant**

Walsh, 2010[e]								
Seasonal Word Comprehension Game	Questioning practices that were not vocabulary-eliciting vs. comparison	29	8.15 (1.61)	8.00 (1.07)	0.15	0.11	+4	> .05
Seasonal Word Production Game	Questioning practices that were not vocabulary-eliciting vs. comparison	29	2.07 (1.69)	1.80 (1.82)	0.27	0.15	+6	> .05
Seasonal Word Comprehension Game	Questioning practices that were vocabulary-eliciting vs. comparison	30	7.40 (2.19)	8.00 (1.07)	–0.60	–0.34	–13	> .05
Seasonal Word Production Game	Questioning practices that were vocabulary-eliciting vs. comparison	30	2.53 (1.19)	1.80 (1.82)	0.73	0.46	+18	> .05
Domain average for comprehension (Walsh, 2010)						**0.10**	**+4**	**Not statistically significant**
Domain average for comprehension (Questioning practices that were not vocabulary-eliciting vs. comparison; Walsh, 2010)						**0.13**	**+5**	**Not statistically significant**
Domain average for comprehension (Questioning practices that were vocabulary-eliciting vs. comparison; Walsh, 2010)						**0.06**	**+3**	**Not statistically significant**
Domain average for comprehension across all studies						**0.20**	**+8**	**na**

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

[a] For Bochna (2010), a difference-in-differences adjustment was needed. The WWC calculated the program group mean using a difference-in-differences approach (see WWC Procedures and Standards Handbook) by adding the impact of the program (i.e., difference in mean gains between the intervention and comparison groups) to the unadjusted comparison group posttest means. This study is characterized as having a statistically significant positive effect because the mean effect is positive and statistically significant. Please see the WWC Procedures and Standards Handbook (version 3.0) for more information.

[b] Box & Aldridge (1993) did not report posttest standard deviations; therefore, the WWC could not calculate effect sizes and improvement indices, and the study does not contribute to the effectiveness rating. The authors present ANCOVA findings across all three conditions indicating that there are statistically significant differences between groups at posttest, but they do not report statistical significance for separate tests of *shared book reading* vs. each of the comparison conditions.

[c] For Lamb (1986), three groups are included in the review—one intervention group and two comparison groups. Both contrasts are of interest to the review. A correction for multiple comparisons was needed. The multiple comparison correction did not change the significance of findings reported by the author. Findings are based on unadjusted posttest means and standard deviations. The author also presented ANCOVA findings that adjusted for pretest; however, they did not test the effects of the conditions of interest for this report. This study is characterized as having an indeterminate effect because the mean effect is neither statistically significant nor substantively important. Please see the WWC Procedures and Standards Handbook (version 3.0) for more information.

[d] For Pollard-Durodola et al. (2011), a correction for multiple comparisons was needed but did not affect whether any of the contrasts were found to be statistically significant. The *p*-values presented here were reported in the original study. The WWC calculated the program group mean using a difference-in-differences approach (see WWC Procedures and Standards Handbook) by adding the impact of the program (i.e., difference in mean gains between the intervention and comparison groups) to the unadjusted comparison group posttest means. This study is characterized as having an indeterminate effect because the mean effect is neither statistically significant nor substantively important. Please see the WWC Procedures and Standards Handbook (version 3.0) for more information.

[e] Walsh (2010) includes three groups—two intervention groups and one comparison group. Both contrasts are of interest to the review. A correction for multiple comparisons was needed. The multiple comparison correction did not change the significance of findings reported by the author. A difference-in-differences adjustment was needed for the Seasonal Word Comprehension Game outcome; the WWC calculated the program group mean for Word Comprehension using a difference-in-differences approach (see WWC Procedures and Standards Handbook) by adding the impact of the program (i.e., difference in mean gains between the intervention and comparison groups) to the unadjusted comparison group posttest means. This study is characterized as having an indeterminate effect because the mean effect is neither statistically significant nor substantively important. Please see the WWC Procedures and Standards Handbook (version 3.0) for more information.

Appendix C.2: Findings included in the rating for the language development domain

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			*p*-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
Justice et al., 2010[a]								
Composite Measure of Language Ability	High dose shared book reading vs. read aloud	278	91.67 (13.91)	90.61 (13.81)	1.06	0.08	+3	.65
Domain average for language development (Justice et al., 2010)						**0.08**	**+3**	**Not statistically significant**
Lamb, 1986[b]								
Record of Oral Language	Reading & interaction vs. reading without interaction	19	8.75 (8.30)	13.88 (10.51)	–5.13	–0.52	–20	> .05
Record of Oral Language	Reading & interaction vs. interaction without reading	19	8.75 (8.30)	9.44 (7.14)	–0.69	–0.08	–3	> .05
Domain average for language development (Lamb, 1986)						**–0.30**	**–12**	**Not statistically significant**
Domain average for language development (reading & interaction vs. reading without interaction; Lamb, 1986)						**–0.52**	**–20**	**Not statistically significant**
Domain average for language development (reading & interaction vs. interaction without reading; Lamb, 1986)						**–0.08**	**–3**	**Not statistically significant**
Mason et al., 1990[c]								
Test of Early Language Development (TELD)	4-year-old preschool children	232	na	na	na	0.07	+3	> .05
Domain average for language development (Mason et al., 1990)						**0.07**	**+3**	**Not statistically significant**
Mautte, 1991[d]								
Preschool Language Scale	4-year-old preschool children	38	65.01 (5.62)	61.30 (9.34)	3.71	0.47	+18	> .05
Domain average for language development (Mautte, 1991)						**0.47**	**+18**	**Not statistically significant**
Domain average for language development across all studies						**0.08**	**+3**	**na**

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

[a] For Justice et al. (2010), a correction for clustering was needed but did not affect whether the contrast was found to be statistically significant. The authors present impact estimates from hierarchical linear models (HLMs) controlling for post-intervention classroom quality; therefore, the WWC does not present the author-reported estimates as they may be biased due to the inclusion of an endogenous covariate in the analysis. The WWC calculated the program group mean using a difference-in-differences approach (see WWC Procedures and Standards Handbook) by adding the impact of the program (i.e., difference in mean gains between the intervention and comparison groups) to the unadjusted comparison group post-test means. This study is characterized as having an indeterminate effect because the mean effect is neither statistically significant nor substantively important. The *p*-value presented here was reported in the original study. Please see the WWC Procedures and Standards Handbook (version 3.0) for more information.

[b] For Lamb (1986), three groups are included in the review—one intervention group and two comparison groups. Both contrasts are of interest to the review. A correction for multiple comparisons was needed. The multiple comparison correction did not change the significance of findings reported by the author. Findings are based on unadjusted posttest means and standard deviations. The author also presented ANCOVA findings that adjusted for pretest; however, they did not test the effects of the conditions of interest for this report. This study is characterized as having a substantively important negative effect as the mean effect is negative and larger than 0.25 SD but is not statistically significant. Please see the WWC Procedures and Standards Handbook (version 3.0) for more information.

[c] For Mason et al. (1990), a correction for clustering was needed but did not affect whether any of the contrasts were found to be statistically significant. The authors reported that the results of an ANOVA that found no significant difference between intervention and comparison children on the Test of Early Language Development. This study is characterized as having an indeterminate effect because the mean effect is neither statistically significant nor substantively important. Please see the WWC Procedures and Standards Handbook (version 3.0) for more information.

[d] For Mautte (1991), a difference-in-differences adjustment was needed. The author presented posttest means that are adjusted across all three conditions in the study; however, only two of the groups are of interest to this review. Therefore, the WWC calculated the program group mean using a difference-in-differences approach (see WWC Procedures and Standards Handbook) by adding the impact of the program (i.e., difference in mean gains between the intervention and comparison groups) to the unadjusted comparison group post-test means. The *p*-value presented here was reported in the original study. This study is characterized as having a substantively important positive effect because the mean effect is positive and larger than 0.25 SD but is not statistically significant. Please see the WWC Procedures and Standards Handbook (version 3.0) for more information.

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

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			*p*-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
Box & Aldridge, 1993[a]								
Early School Inventory–Preliteracy/Part A: Print Concepts Subtest	Shared reading vs. usual instruction	48	15.20 (nr)	10.70 (nr)	4.50	na	na	na
Early School Inventory–Preliteracy/Part A: Print Concepts Subtest	Shared reading vs. movement activities	49	15.20 (nr)	10.76 (nr)	4.44	na	na	na
Domain average for alphabetics (Box & Aldridge, 1993)						**na**	**na**	**na**
Justice et al., 2010[b]								
Composite Measure of Print Knowledge	High dose shared book reading vs. read aloud	288	0.09 (0.83)	–0.11 (0.88)	0.20	0.23	+9	< .05
Domain average for alphabetics (Justice et al., 2010)						**0.23**	**+9**	**Not statistically significant**
Lamb, 1986[c]								
Concepts About Print Test	Reading & interaction vs. reading without interaction	19	3.60 (2.95)	4.22 (2.05)	–0.62	–0.23	–9	**> .05**
Concepts About Print Test	Reading & interaction vs. interaction without reading	19	3.60 (2.95)	2.33 (2.74)	1.27	0.43	+17	**> .05**
Domain average for alphabetics (Lamb, 1986)						**0.10**	**+4**	**Not statistically significant**

Domain average for alphabetics (reading & interaction vs. reading without interaction; Lamb, 1986)	**–0.23**	**–9**	**Not statistically significant**
Domain average for alphabetics (reading & interaction vs. interaction without reading; Lamb, 1986)	**0.43**	**+17**	**Not statistically significant**
Domain average for alphabetics across all studies	**0.13**	**+5**	**na**

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

[a] Box & Aldridge (1993) did not report posttest standard deviations; therefore, the WWC could not calculate effect sizes and improvement indices, and the study does not contribute to the effectiveness rating. The authors do present ANCOVA findings across all three conditions indicating that there are statistically significant differences between groups at posttest, but they do not report statistical significance for separate tests of *shared book reading* vs. each of the comparison conditions.

[b] For Justice et al. (2010), the *p*-value presented here was reported in the original study. A correction for clustering was needed and resulted in a WWC-computed *p*-value of .14 for the Composite Measure of Print Knowledge; therefore, the WWC does not find the result to be statistically significant. The authors presented impact estimates from HLMs controlling for post-intervention classroom quality; therefore, the WWC does not present the author-reported estimates as they may be biased due to the inclusion of an endogenous covariate in the analysis. The WWC calculated the program group mean using a difference-in-differences approach (see WWC Procedures and Standards Handbook) by adding the impact of the program (i.e., difference in mean gains between the intervention and comparison groups) to the unadjusted comparison group posttest means. This study is characterized as having an indeterminate effect because the mean effect is neither statistically significant nor substantively important. Please see the WWC Procedures and Standards Handbook (version 3.0) for more information.

[c] For Lamb (1986), three groups are included in the review—one intervention group and two comparison groups. Both contrasts are of interest for the review. A correction for multiple comparisons was needed. The multiple comparison correction did not change the significance of findings reported by the author. Findings are based on unadjusted posttest means and standard deviations. The author also presented ANCOVA findings that adjusted for pretest; however, they did not test the effects of the conditions of interest for this report. This study is characterized as having an indeterminate effect because the mean effect is neither statistically significant nor substantively important. Please see the WWC Procedures and Standards Handbook (version 3.0) for more information.

Appendix C.4: Findings included in the rating for the general reading achievement domain

			Mean (standard deviation)		WWC calculations			
Outcome measure	**Study sample**	**Sample size**	**Intervention group**	**Comparison group**	**Mean difference**	**Effect size**	**Improvement index**	***p*-value**
Mautte, 1991[a]								
Test of Early Reading Ability (TERA)	4-year-old preschool children	38	8.47 (4.69)	9.90 (8.17)	–1.43	–0.21	–8	> .05
Domain average for general reading achievement (Mautte, 1991)						**–0.21**	**–8**	**Not statistically significant**
Domain average for general reading achievement across all studies						**–0.21**	**–8**	**na**

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

[a] For Mautte (1991), a difference-in-differences adjustment was needed. The author presented posttest means that are adjusted across all three conditions in the study; however, only two of the groups are of interest to this review. Therefore, the WWC calculated the program group mean using a difference-in-differences approach (see WWC Procedures and Standards Handbook) by adding the impact of the program (i.e., difference in mean gains between the intervention and comparison groups) to the unadjusted comparison group posttest means. The *p*-value presented here was reported in the original study. This study is characterized as having an indeterminate effect because the mean effect is neither statistically significant nor substantively important. Please see the WWC Procedures and Standards Handbook (version 3.0) for more information.

Appendix D.1: Supplemental follow-up findings for the comprehension domain

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			*p*-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
Piasta, 2012 high-dose (4 sessions/week), 1-year follow-up[a]								
Woodcock-Johnson III (WJ-III) Passage Comprehension Subtest	High dose vs. comparison	244	9.09 (3.74)	8.11 (2.91)	0.98	0.29	11	< .01
Peabody Picture Vocabulary Test–4 (PPVT-4)	High dose vs. comparison	246	96.22 (10.81)	97.84 (13.70)	–1.62	–0.13	–5	> .05
Piasta, 2012 low-dose (2 sessions/week), 1-year follow-up[a]								
WJ-III Passage Comprehension Subtest	Low dose vs. comparison	218	8.36 (408)	8.11 (2.91)	0.25	0.07	+3	.60
PPVT-4	Low dose vs. comparison	221	97.22 (10.50)	97.84 (13.70)	–0.62	–0.05	–2	> .05
Walsh, 2010 follow-up after 1 week[b]								
Seasonal Word Comprehension Game	Questioning practices that were not vocabulary-eliciting vs. comparison	29	7.58 (2.85)	7.60 (2.61)	–0.02	–0.01	0	> .05
Seasonal Word Production Game	Questioning practices that were not vocabulary-eliciting vs. comparison	29	2.29 (2.52)	1.73 (1.75)	0.56	0.25	+10	> .05
Seasonal Word Comprehension Game	Questioning practices that were vocabulary-eliciting vs. comparison	30	7.14 (2.80)	7.60 (2.61)	–0.46	–0.17	–7	> .05
Seasonal Word Production Game	Questioning practices that were vocabulary-eliciting vs. comparison	30	2.00 (1.41)	1.73 (1.75)	0.27	0.17	+7	> .05

Table Notes: The supplemental findings presented in this table are additional findings from studies in this report that do not factor into the determination of the intervention rating because they are delayed effects rather than findings from the end of the intervention. For mean difference, effect size, and improvement index values reported in the table, a positive number favors the intervention group and a negative number favors the comparison group. The effect size is a standardized measure of the effect of an intervention on outcomes, representing the average change expected for all individuals who are given the intervention (measured in standard deviations of the outcome measure). The improvement index is an alternate presentation of the effect size, reflecting the change in an average individual's percentile rank that can be expected if the individual is given the intervention.

[a] Piasta (2012) presents follow-up, supplementary results for Justice et al. (2010). A correction for multiple comparisons was needed but did not affect whether any of the contrasts were found to be statistically significant. The author presented impact estimates from HLMs controlling for baseline measures. The program group mean is calculated by adding the impact of the program (i.e., difference in mean gains between the intervention and comparison groups) to the unadjusted comparison group posttest mean. The effect sizes presented here were calculated by the WWC using the Hedges' *g* formula; the author-reported effect sizes used the Cohen's *d* formula. The *p*-values presented here were reported in the original study.

[b] Walsh (2010) includes three groups—two intervention groups and one comparison group. Both contrasts are of interest to the review. A correction for multiple comparisons was needed but did not affect whether any of the contrasts were found to be statistically significant. The *p*-values presented here were reported in the original study. The WWC calculated the program group mean for the Seasonal Word Comprehension Game outcome using a difference-in-differences approach (see WWC Procedures and Standards Handbook) by adding the impact of the program (i.e., difference in mean gains between the intervention and comparison groups) to the unadjusted comparison group posttest means. Please see the WWC Procedures and Standards Handbook (version 3.0) for more information.

Appendix D.2: Supplemental subgroup findings for the language development domain

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			*p*-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
Mautte, 1991 delayed development subgroup (immediate posttest)[a]								
Preschool Language Scale	Developmentally delayed 4-year-old preschool children	21	60.10 (3.60)	54.45 (6.20)	5.65	1.07	+36	nr

Table Notes: The supplemental findings presented in this table are additional findings from studies in this report that do not factor into the determination of the intervention rating because they are effects only for a subgroup of children. For mean difference, effect size, and improvement index values reported in the table, a positive number favors the intervention group and a negative number favors the comparison group. The effect size is a standardized measure of the effect of an intervention on outcomes, representing the average change expected for all individuals who are given the intervention (measured in standard deviations of the outcome measure). The improvement index is an alternate presentation of the effect size, reflecting the change in an average individual's percentile rank that can be expected if the individual is given the intervention. Some statistics may not sum as expected due to rounding. nr = not reported.

[a] Mautte (1991) presents ANOVA findings indicating that the three groups vary depending on developmental level. However, the author did not report the *p*-value for an analysis comparing shared book reading and the reading-only comparison group for the developmentally delayed subgroup.

Appendix D.3: Supplemental 1-year and 2-year follow-up findings for the alphabetics domain

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			*p*-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
Piasta, 2012 low-dose (2 sessions/week), 1-year follow-up[a]								
Woodcock-Johnson III (WJ-III) Letter–Word Identification Subtest	Low dose vs. comparison	221	20.76 (7.17)	19.32 (5.24)	1.44	0.23	+9	.06
WJ-III Spelling Subtest	Low dose vs. comparison	221	15.30 (3.29)	14.58 (2.95)	0.72	0.23	+9	.06
Piasta, 2012 high-dose (4 sessions/week), 2-year follow-up[a]								
WJ-III Letter–Word Identification Subtest	High dose vs. comparison	250	33.83 (8.94)	31.21 (8.45)	2.62	0.30	+12	.02
WJ-III Spelling Subtest	High dose vs. comparison	250	21.82 (5.02)	19.99 (4.62)	1.83	0.38	+15	< .01
Piasta, 2012 low-dose (2 sessions/week), 2-year follow-up[a]								
WJ-III Letter–Word Identification Subtest	Low dose vs. comparison	234	31.81 (8.30)	31.21 (8.45)	0.60	0.07	+3	.58
WJ-III Spelling Subtest	Low dose vs. comparison	236	20.02 (4.76)	19.99 (4.62)	1.03	0.22	+9	< .05

Table Notes: The supplemental findings presented in this table are additional findings from studies in this report that do not factor into the determination of the intervention rating because they are delayed effects rather than findings from the end of the intervention. For mean difference, effect size, and improvement index values reported in the table, a positive number favors the intervention group and a negative number favors the comparison group. The effect size is a standardized measure of the effect of an intervention on outcomes, representing the average change expected for all individuals who are given the intervention (measured in standard deviations of the outcome measure). The improvement index is an alternate presentation of the effect size, reflecting the change in an average individual's percentile rank that can be expected if the individual is given the intervention. Some statistics may not sum as expected due to rounding.

[a] Piasta (2012) presents follow-up, supplementary results for Justice et al. (2010). A correction for multiple comparisons was needed but did not affect whether any of the contrasts were found to be statistically significant. The author presented impact estimates from HLMs controlling for baseline measures. The program group mean is calculated by adding the impact of the program (i.e., difference in mean gains between the intervention and comparison groups) to the unadjusted comparison group posttest mean. The effect sizes presented here were calculated by the WWC using the Hedges' *g* formula; the author-reported effect sizes used the Cohen's *d* formula. The *p*-values presented here were reported in the original study.

Appendix D.4: Supplemental subgroup findings for the general reading achievement domain

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			*p*-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
Mautte, 1991 delayed development subgroup (immediate posttest)[a]								
Test of Early Reading Ability (TERA)	Developmentally delayed 4-year-old preschool children	21	7.00 (2.79)	6.00 (2.68)	1.00	0.35	+14	nr

Table Notes: The supplemental findings presented in this table are additional findings from studies in this report that do not factor into the determination of the intervention rating because they are effects only for a subgroup of children. For mean difference, effect size, and improvement index values reported in the table, a positive number favors the intervention group and a negative number favors the comparison group. The effect size is a standardized measure of the effect of an intervention on outcomes, representing the average change expected for all individuals who are given the intervention (measured in standard deviations of the outcome measure). The improvement index is an alternate presentation of the effect size, reflecting the change in an average individual's percentile rank that can be expected if the individual is given the intervention. Some statistics may not sum as expected due to rounding. nr = not reported.

[a] Mautte (1991) presents two-way ANOVA findings for the three groups vary by developmental level. However, the author did not report a *p*-value for an analysis comparing *shared book reading* and the reading-only comparison group for the developmentally delayed subgroup.

Endnotes

[1] The descriptive information for this program was obtained from publicly available sources: the previous What Works Clearinghouse intervention reports on *shared book reading* (September 2006) and *interactive shared book reading* (January 2007). According to WWC procedures, the review team leadership (which includes a lead methodologist and content area expert) can make determinations within the review protocol about which practices to group together for intervention reports. In this case, the revew team leadership reviewed three previously released intervention reports (*shared book reading*, *interactive shared book reading*, and *dialogic reading*) and determined that the *shared book reading* and *interactive shared book reading* intervention reports should be combined with updating the review. *Dialogic reading* was determined to be a distinct practice and appears as a separate intervention report. The WWC requests developers review the program description sections for accuracy from their perspective. However, because *shared book reading* is a practice that does not have a single developer responsible for providing information and materials, there was no developer available to review the program description. Further verification of the accuracy of the descriptive information for this program is beyond the scope of this review.

[2] The WWC intervention report on *dialogic reading* can be accessed from the What Works Clearinghouse website (http://ies.ed.gov/ncee/wwc/).

[3] The literature search reflects documents publicly available by July 2014. The literature search was conducted as described in the Early Childhood Education protocol, version 3.0. The WWC developed a definition of *shared book reading* (detailed in the program description) for use in determining whether a study fell within the scope of this review. The previous reports on *shared book reading* and *interactive shared book reading* were released in September 2006 and January 2007, respectively. For this updated intervention report, the Early Childhood Education topic area refined the *shared book reading* program description, which affected the studies and contrasts of interest, compared to the earlier reports. The clarifications of the program description specify that the adults used one or more interactive strategies when reading to children to draw attention to elements of the book or reading passage being read. That is, in order to be considered a test of *shared book reading*, the intervention condition must include both reading aloud activities, *and* the use of one or more planned or structured interactive strategies focused on drawing attention to the text being read, and the comparison condition must exclude one or both of these key components. Therefore, conditions that involve an adult reading a book to children straight through without interaction are no longer considered *shared book reading*, and instead are considered an appropriate comparison condition. Similarly, conditions that involve interaction activities (in particular, unstructured interaction activities that do not draw attention to the text being read), but do not include any read-aloud components are plausible comparison conditions. Furthermore, studies that compare two book reading conditions that both involve interactive strategies are not considered a test of the effect of *shared book reading*, although they may have been for a previous report. Finally, studies that compare *shared book reading* interventions in preschool classrooms to preschool teachers' usual reading style (i.e., business-as-usual read-aloud) or a no-reading condition are both considered a test of the effect of *shared book reading* compared to business-as-usual classroom experiences.

This report has been updated to include reviews of 38 studies that have been released since 2006. Of the additional studies, 16 were not within the scope of the review protocol for the Early Childhood Education topic area, 12 did not use an eligible design, and five did not use a sample aligned with the protocol. There was one new study that met screening criteria for the Early Childhood Education topic area review protocol but did not meet WWC group design standards. Four new studies met WWC group design standards (with or without reservations): Bochna (2010); Justice et al. (2010); Pollard-Durodola et al. (2011); and Walsh (2010). A complete list and disposition of all studies reviewed are provided in the references.

This report includes reviews of all 15 studies in the previous reports on *shared book reading* and *interactive shared book reading*. The report confirms the study rating for five studies: Lamb (1986) and Mautte (1991) met WWC evidence standards without reservations in the previous report, and Kertoy (1994), Morrow (1998), and Murray et al. (1996) did not meet WWC evidence standards in a previous report. This report resulted in revised dispositions for 10 studies.

Two studies, Box & Aldridge (1993) and Mason et al. (1990), which did not meet WWC evidence standards in the previous reports, both meet WWC group design standards without reservations using version 3.0 standards, because both studies are randomized controlled trials (RCTs) with low attrition. These revised dispositions are due to revisions to the WWC standards. In particular, in the WWC version 1.0 standards, studies received a rating of *does not meet WWC evidence standards* if the author did not provide sufficient information to calculate an effect size, even if the study design met standards with or without reservations. For Box & Aldridge (1993), insufficient information was provided to calculate effect sizes; therefore, the study does not contribute to the effectiveness rating for the intervention in the current report. For Mason et al. (1990), the WWC was able to calculate an effect size based on the author-reported sample size and ANOVA *F*-statistic.

One study, Irlen (2003), which met WWC evidence standards without reservations in the previous report, *does not meet WWC group design standards* using version 3.0 standards. The revised disposition for this study is due to revisions to the WWC standards. In

particular, requirements for calculating attrition rates and establishing baseline equivalence were different in the WWC version 1.0 standards. The information needed to calculate attrition rates and establish baseline equivalence in accordance with the WWC version 3.0 standards is not publicly available.

Five studies that were included in the previous intervention reports—two that met WWC evidence standards without reservations, one that met WWC evidence standards with reservations, and two that did not meet WWC evidence standards—were determined not to be within the scope of the review protocol for Early Childhood Education (version 3.0). The revised dispositions for these five studies are due to clarifications in the definition of *shared book reading* used for the updated intervention report. The revised definition indicates that the adults used one or more planned or structured interactive strategies to draw attention to elements of the book or reading passage being read. For this updated intervention report, studies in which all conditions involve interactive strategies during shared reading are not considered a test the effect of *shared book reading*. Based on the updated definition, the five studies do not test the effect of *shared book reading*. In three of the studies—one that previously met WWC evidence standards without reservations (Justice & Ezell, 2002); one that previously met WWC evidence standards with reservations (McCormick & Mason, 1989); and one that previously did not meet WWC evidence standards (Reese & Cox, 1999)—all conditions were variations of *shared book reading*, without a non-*shared book reading* comparison condition. In two studies—one that previously met WWC evidence standards (Lonigan et al., 1999) and one that previously did not meet WWC evidence standards (Ratliff, 1987), none of the conditions were *shared book reading*, per se. Lonigan et al. (1999) included dialogic reading, which is a specific type of *shared book reading* addressed in a separate intervention report. Ratliff (1987) included *shared book reading* bundled with other activities so that the observed impact estimate confounds the effect of *shared book reading* with the other activities.

One study, Jones (1987), which did not meet WWC evidence standards in the previous report, was determined to have a sample that was not aligned with the Early Childhood Education review protocol (version 3.0), because more than half of the children in the sample had disabilities. In addition, the study had a cohort comparison design, which is not an eligible design, as stated in the WWC Procedures and Standards Handbook (version 3.0).

One last study, Neuman (1996), which did not meet WWC evidence standards in the previous report, was determined not to have an eligible design, because the study did not have a comparison group. This revised disposition is due to revisions to the WWC standards. In particular, in the WWC version 1.0 standards, studies received a rating of *does not meet WWC standards* if the design was not eligible; however, in the version 3.0 standards, studies without an eligible design are not eligible for review.

The studies in this report were reviewed using the Standards from the WWC Procedures and Standards Handbook (version 3.0), along with those described in the Early Childhood Education review protocol (version 3.0). The evidence presented in this report is based on available research. Findings and conclusions may change as new research becomes available.

[4] The other four Early Childhood Education domains are cognition, mathematics, social-emotional development, and fluency.

[5] For criteria used in the determination of the rating of effectiveness and extent of evidence, see the WWC Rating Criteria on p. 41. These improvement index numbers show the average and range of student-level improvement indices for all findings across the studies.

[6] For Box & Aldridge (1993), the WWC was unable to calculate effect sizes or statistical significance for effects on alphabetics and comprehension, as posttest standard deviations were not available; thus, the study does not contribute to the effectiveness rating. In the previous *interactive shared book reading* intervention report (January 2007), this study received a rating of *does not meet WWC evidence standards* following the WWC Procedures and Standards Handbook (version 1.0), despite being a randomized controlled trial with low attrition, because the information to calculate effect sizes was not reported.

[7] In the previous *interactive shared book reading* intervention report (January 2007), for Lamb (1986), the intervention—shared reading with structured language interaction—was compared to only the read-aloud without interaction comparison. However, in the current intervention report, the interaction without read-aloud condition is also considered an acceptable comparison for testing the effect of *shared book reading*.

[8] The author (Lamb, 1986) implemented all three intervention conditions but had no or minimal contact with children in the no-intervention comparison group; therefore, the agent of the intervention was confounded with the intervention for comparisons between the read-aloud with structured language interaction and the no-intervention comparison group (i.e., the effects of the person providing the intervention cannot be separated from the effects of the intervention).

[9] The authors (Mason et al., 1990) also assessed two outcomes in the alphabetics domain and one outcome in the general reading achievement domain; however, the outcomes did not meet WWC requirements for reliability and were excluded from this report. Additional details on the measures are provided in Endnote 15.

[10] The author (Mautte, 1991) implemented both intervention conditions but had no or minimal contact with children in the no-intervention, business-as-usual general classroom instruction comparison group; therefore, the agent of the intervention was confounded with

the intervention for comparisons between the repeated reading plus adult interaction and the no-intervention comparison group (i.e., the effects of the person providing the intervention cannot be separated from the effects of the intervention).

[11] Follow-up results 1 year and 2 years post-intervention are reported in Piasta et al. (2012) and summarized in Appendix D. Follow-up results are reported for both the high-dose and low-dose shared reading intervention groups compared to the business-as-usual read-aloud comparison group. No studies report immediate posttest results for the low-dose shared reading intervention condition vs. the business-as-usual read-aloud comparison condition.

[12] The authors (Box & Aldridge, 1993) present ANCOVA findings across all three conditions (*shared book reading* and the two comparison conditions), indicating that there are statistically significant differences at posttest, but they do not report statistical significance for tests of *shared book reading* vs. each of the comparison conditions.

[13] The authors (Box & Aldridge, 1993) assessed another outcome in the alphabetics domain using the Concepts About Print Test (Clay, 1979); however, the Concepts About Print Test did not meet requirements for reliability and was excluded from this report.

[14] In the previous *interactive shared book reading* intervention report (January 2007), for Lamb (1986), the intervention—shared reading with structured language interaction—was compared to only the read-aloud without interaction comparison. However, in the current intervention report, the interaction without read-aloud condition is also considered an acceptable comparison for testing the effect of *shared book reading*.

[15] The authors (Mason et al., 1990) also assessed two outcomes in the alphabetics domain and one outcome in the general reading achievement domain. All three outcomes were subscales of the Beginning Educational Assessment–3: Print Concepts, Letter Knowledge, and Beginning Reading. The outcomes were excluded from this report because they did not meet WWC outcome requirements for reliability. It could not be confirmed that the Beginning Educational Assessment–3 was a standardized measure, and reliability information was not reported.

[16] The author (Mautte, 1991) implemented both intervention conditions but had no or minimal contact with children in the no-intervention, business-as-usual general classroom instruction comparison group; therefore, the agent of the intervention was confounded with the intervention for comparisons between the repeated reading plus adult interaction and the no-intervention comparison group (i.e., the effects of the person providing the intervention cannot be separated from the effects of the intervention).

[17] The WWC identified two additional sources related to Justice et al, 2010. One of these studies (Justice et al, 2009, which reports interim findings) does not contribute unique information to Appendix A.7 and is not listed here; the other (Piasta et al, 2012, which reports 1-year and 2-year follow-up findings) does contribute unique information to Appendix A.7 and is listed.

[18] The authors (Pollard-Durodola et al., 2011) assessed two additional measures in the comprehension domain. Both were researcher-developed measures: the Receptive Picture Vocabulary Test and the Expressive Picture Vocabulary Test, which were excluded from this report. The outcomes did not meet WWC requirements because they were overaligned with the intervention condition.

Recommended Citation

U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse. (2015, April). *Early Childhood Education intervention report: Shared book reading.* Retrieved from http://whatworks.ed.gov

WWC Rating Criteria

Criteria used to determine the rating of a study

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

Criteria used to determine the rating of effectiveness for an intervention

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

Criteria used to determine the extent of evidence for an intervention

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

Glossary of Terms

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

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

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

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

Domain A domain is a group of closely related outcomes.

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

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

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

Extent of evidence An indication of how much evidence supports the findings. The criteria for the extent of evidence levels are given in the WWC Rating Criteria on p. 41.

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

Intervention An educational program, product, practice, or policy aimed at improving student outcomes.

Intervention report A summary of the findings of the highest-quality research on a given program, product, practice, or policy in education. The WWC searches for all research studies on an intervention, reviews each against design standards, and summarizes the findings of those that meet WWC design standards.

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

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

Randomized controlled trial (RCT) A randomized controlled trial (RCT) is an experiment in which eligible study participants are randomly assigned to intervention and comparison groups.

Rating of effectiveness The WWC rates the effects of an intervention in each domain based on the quality of the research design and the magnitude, statistical significance, and consistency in findings. The criteria for the ratings of effectiveness are given in the WWC Rating Criteria on p. 41.

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

Glossary of Terms

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

Statistical significance Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups. The WWC labels a finding statistically significant if the likelihood that the difference is due to chance is less than 5% ($p < .05$).

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

Systematic review A review of existing literature on a topic that is identified and reviewed using explicit methods. A WWC systematic review has five steps: 1) developing a review protocol; 2) searching the literature; 3) reviewing studies, including screening studies for eligibility, reviewing the methodological quality of each study, and reporting on high quality studies and their findings; 4) combining findings within and across studies; and, 5) summarizing the review.

Please see the WWC Procedures and Standards Handbook (version 3.0) for additional details.