

Appendix

Appendix A1.1 Study characteristics: Crain-Thoreson & Dale, 1999¹ (randomized controlled trial)

Characteristic	Description
Study citation	Crain-Thoreson, C., & Dale, P. S. (1999). Enhancing linguistic performance: Parents and teachers as book reading partners for children with language delays. <i>Topics in Early Childhood Special Education, 19</i> (1), 28–39.
Participants	The study began with 37 children. All children were eligible for early childhood special-education services and had mild to moderate language delays. Based on pretest scores, triads of children with similar receptive vocabulary scores were formed within each district. Within each triad, one member was randomly assigned to a parent dialogic reading group, one member was assigned to a staff dialogic reading group, and one was assigned to a control condition. Thirteen children were assigned to the staff-implemented group, which was not included in this review, and five more children did not complete the study, leaving 19 children remaining in the sample. The mean age of all the children who completed the study was 51.6 months (ranging from 39 to 66 months), and 31.3% of these children were female. ² The mean age of children in the sample included in this review was 50.5 months.
Setting	The study took place in five classrooms in five schools in three school districts in the Pacific Northwest. Children attended publicly funded preschool programs that provided early intervention for children with special needs.
Intervention	The study included two intervention groups: one in which program staff implemented dialogic reading, and another in which parents implemented dialogic reading. The comparison between the staff-implemented group and the no-treatment comparison group did not meet evidence standards because of high differential attrition and lack of baseline equivalence and is excluded from this report. Dialogic reading was implemented over an eight-week period, during which staff or parents engaged in book reading with individual children at least four times per week.
Comparison	Children in the control group did not participate in dialogic reading. They participated in group story time, which was the standard practice.
Primary outcomes and measurement	The primary outcome domain was children's communication and language competencies, measured by three nonstandardized measures and two standardized measures. The nonstandardized measures included mean length of utterances, number of utterances, and number of different words used (lexical diversity). Children's vocabulary knowledge was measured by two standardized tests: the Peabody Picture Vocabulary Test–Revised and the Expressive One-Word Picture Vocabulary Test–Revised. For a more detailed description of these outcome measures, see Appendix A2.
Staff/teacher/parent training	Parents were trained in the dialogic reading program during two 90-minute instructional sessions held four weeks apart. Videotape training, live demonstration, and role-play were used during the training. Handouts summarizing the training components were provided to parents. The researchers modified the parent training program to address the needs of students with language delays by teaching parents to pause and give their children time to respond.

1. The Crain-Thoreson and Dale (1999) study was previously reviewed for the WWC's intervention report on dialogic reading conducted as part of the Early Childhood Education (ECE) review of interventions for the general population. The parent-implemented intervention group was excluded from the previous report, because the ECE review focused exclusively on center-based interventions. The Early Childhood Education Interventions for Children with Disabilities topic area includes both center-based and home-based interventions in the review. For the current report, we report only the comparison between the parent-implemented intervention group and the no-treatment comparison group. The comparison between the staff-implemented intervention group and the no-treatment control group did not meet evidence standards because of high differential attrition and lack of baseline equivalence and is excluded from this report.
2. Crain-Thoreson and Dale (1999) report the mean age of children in all three groups (staff-implemented dialogic reading, parent-implemented dialogic reading, and no-treatment control) who completed the study. The comparison between the staff-implemented group and the no-treatment comparison group did not meet evidence standards because of high differential attrition and lack of baseline equivalence and is excluded from this report.

Appendix A1.2 Study characteristics: Dale, Crain-Thoreson, Notari-Syverson, & Cole, 1996 (randomized controlled trial)

Characteristic	Description
Study citation	Dale, P. S., Crain-Thoreson, C., Notari-Syverson, A., & Cole, K. (1996). Parent-child book reading as an intervention technique for young children with language delays. <i>Topics in Early Childhood Special Education, 16</i> (2), 213–235.
Participants	Thirty-three mother-child dyads were recruited over a two-year period from an early childhood education center. The children were ages three to six and had mild to moderate language delays, functioning at the two- to four-year-old level; 27% of the sample was female. Average baseline test scores on the Peabody Picture Vocabulary Test–Revised were more than 2 standard deviations below the normed mean. English was the primary language for all dyads. The dyads were assigned randomly either to dialogic reading or to the comparison condition, the <i>Conversational Language Training Program</i> .
Setting	Study participants were recruited from programs at the Experimental Education Unit at the University of Washington. Parent training sessions were conducted in small groups in therapy/lab rooms.
Intervention	Parents attended two training sessions with their children, where they viewed a videotaped presentation of effective ways of facilitating language development through reading and then had a brief group discussion. Modules helped facilitate appropriate parental feedback to children during book reading. The intervention included handouts that summarized each of the modules and included examples. Parents were asked to implement the intervention at home over a six- to eight-week period. Books and handouts were given to parents to take home with them. Researchers were unable to assess the consistency of parental implementation of this program.
Comparison	The comparison program, the <i>Conversational Language Training Program</i> , was designed to be comparable both in presentation and in general nature to the intervention but did not include books. Both the intervention and the comparison program emphasized an interactive, responsive style of communication with children. Parents in the comparison group attended two training sessions in which they viewed videotapes, received handouts, were given one of two toys, and were asked to implement the comparison program at home over a six- to eight-week period.
Primary outcomes and measurement	The primary outcome domain was children's communication and language competencies. The study used three nonstandardized measures of communication and language competencies: mean length utterance, number of different words used (lexical diversity), and total number of utterances. The study used these measures for two types of activities: book reading and play. For a more detailed description of these outcome measures, see Appendix A2.
Staff/teacher/parent training	Parents were trained in dialogic reading in two sessions. Training included a videotaped presentation, group discussion, and handouts summarizing program components.

Appendix A2 Outcome measures for the communication and language competencies domain

Outcome measure ¹	Description
Peabody Picture Vocabulary Test–Revised (PPVT-R)	A standardized measure of children’s receptive vocabulary that requires them to identify pictures that correspond to spoken words (as cited in Crain-Thoreson & Dale, 1999).
Expressive One-Word Picture Vocabulary Test–Revised (EOWPVT-R)	A standardized measure of children’s expressive vocabulary that requires them to name pictures of common objects, actions, and concepts (as cited in Crain-Thoreson & Dale, 1999).
Mean length of utterances (MLU)	The number of child utterances during videotaped activity measured using the Computerized Language Analysis Programs (CLAN) (as cited by Crain-Thoreson & Dale, 1999; Dale et al., 1996).
Number of child utterances	The number of child utterances during videotaped activity measured using CLAN (as cited by Crain-Thoreson & Dale, 1999; Dale et al., 1996).
Lexical diversity	The number of different words spoken by the child during videotaped activity measured using CLAN (as cited by Crain-Thoreson & Dale, 1999; Dale et al., 1996).

1. Process-oriented outcomes measuring the children’s engagement in the tasks—called “partic” in Crain-Thoreson and Dale (1999) and “verbal engagement” or “total engagement” in Dale et al. (1996)—are excluded from this report.

Appendix A3 Summary of study findings included in the rating for the communication and language competencies domain¹

Outcome measure	Study sample	Sample size (children)	Authors' findings from the study		WWC calculations			
			Mean outcome (standard deviation) ²		Mean difference ⁴ (Dialogic reading – comparison)	Effect size ⁵	Statistical significance ⁶ (at $\alpha = 0.05$)	Improvement index ⁷
			Dialogic reading group ³	Comparison group				
Crain-Thoreson & Dale, 1999 (randomized controlled trial)⁸—Parent-led implementation								
PPVT-R	3–5 year olds	19	61.10 (16.50)	62.00 (16.20)	–0.90	–0.05	ns	–2
EOWPVT-R	3–5 year olds	19	74.80 (12.10)	71.00 (10.60)	3.80	0.32	ns	+12
Mean length utterance	3–5 year olds	19	2.75 (0.72)	2.70 (0.97)	0.05	0.06	ns	+2
Total utterances	3–5 year olds	19	38.70 (10.30)	35.30 (21.30)	3.40	0.20	ns	+8
Lexical diversity	3–5 year olds	19	49.00 (13.00)	48.80 (38.70)	0.20	0.01	ns	0
Average for communication and language competencies (Crain-Thoreson & Dale, 1999)⁹						0.11	ns	+4
Dale et al., 1996 (randomized controlled trial)⁸								
Total utterances—book reading	3–6 year olds	33	63.80 (13.90)	40.30 (13.60)	23.50	1.67	Statistically significant	+45
Mean length utterances—book reading	3–6 year olds	33	2.82 (0.76)	2.75 (0.90)	0.07	0.08	ns	+3
Lexical diversity—book reading	3–6 year olds	33	14.60 (12.30)	14.60 (8.30)	0.00	0.00	ns	0
Total utterances—play	3–6 year olds	33	94.50 (36.20)	92.30 (36.40)	2.20	0.06	ns	+2

(continued)

Appendix A3 Summary of study findings included in the rating for the communication and language competencies domain¹ (continued)

Outcome measure	Study sample	Sample size (children)	Authors' findings from the study		WWC calculations			
			Mean outcome (standard deviation) ²		Mean difference ⁴ (Dialogic reading – comparison)	Effect size ⁵	Statistical significance ⁶ (at $\alpha = 0.05$)	Improvement index ⁷
			Dialogic reading group ³	Comparison group				
Mean length utterances—play	3–6 year olds	33	3.08 (1.09)	2.98 (0.81)	0.10	0.10	ns	+4
Lexical diversity—play	3–6 year olds	33	75.00 (20.60)	54.90 (22.00)	20.10	0.92	Statistically significant	+32
Average for communication and language competencies (Dale et al., 1996)⁹						0.47	ns	+18
Domain average for communication and language competencies across all studies⁹						0.29	ns	+11

ns = not statistically significant

PPVT-R = Peabody Picture Vocabulary Test–Revised

EOWPVT-R = Expressive One-Word Picture Vocabulary Test–Revised

1. This appendix reports findings considered for the effectiveness rating and the average improvement indices for the communication and language competencies domain.
2. The standard deviation across all students in each group shows how dispersed the participants' outcomes are: a smaller standard deviation on a given measure would indicate that participants had more similar outcomes.
3. The dialogic reading group mean is the sum of the unadjusted comparison group mean and the adjusted mean difference, which accounts for pretest. Standard deviations are unadjusted.
4. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group. The mean difference is adjusted for pretest differences between the treatment and comparison groups.
5. For an explanation of the effect size calculation, see WWC Procedures and Standards Handbook, Appendix B.
6. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups.
7. The improvement index represents the difference between the percentile rank of the average student in the intervention condition and that of the average student in the comparison condition. The improvement index can take on values between –50 and +50, with positive numbers denoting favorable results for the intervention group.
8. The level of statistical significance was reported by the study authors or, when necessary, calculated by the WWC to correct for clustering within classrooms or schools and for multiple comparisons. For an explanation about the clustering correction, see the WWC Tutorial on Mismatch. For the formulas the WWC used to calculate the statistical significance, see WWC Procedures and Standards Handbook, Appendix C for clustering and WWC Procedures and Standards Handbook, Appendix D for multiple comparisons. For the two dialogic reading studies summarized here, no corrections for clustering were needed; however, corrections for multiple comparisons were needed, so the significance levels may differ from those reported in the original studies.
9. The WWC-computed average effect sizes for each study and for the domain across studies are simple averages rounded to two decimal places. The average improvement indices are calculated from the average effect sizes.

Appendix A4 Dialogic reading rating for the communication and language competencies domain

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

For the outcome domain of communication and language competencies for children with disabilities, the WWC rated dialogic reading as having potentially positive effects. The remaining ratings (mixed effects, no discernible effects, potentially negative effects, negative effects) were not considered, as dialogic reading was assigned the highest applicable rating.

Rating received

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

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

Met. One study found statistically significant positive effects.

AND

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

Met. No studies showed statistically significant or substantively important negative effects.

Other ratings considered

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

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

Not met. Only one study of dialogic reading showed a statistically significant positive effect.

AND

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

Met. No studies showed statistically significant or substantively important negative effects.

1. For rating purposes, the WWC considers the statistical significance of individual outcomes and the domain-level effect. The WWC also considers the size of the domain-level effect for ratings of potentially positive or potentially negative effects. For a complete description, see the WWC Procedures and Standards Handbook, Appendix E.

Appendix A5 Extent of evidence by domain

Outcome domain	Number of studies	Sample size		Extent of evidence ²
		Schools/centers ¹	Students	
Cognitive development	0	na	na	na
Communication/language competencies	2	6	52	Small
Literacy	0	na	na	na
Math competencies	0	na	na	na
Social-emotional development/behavior	0	na	na	na
Functional abilities	0	na	na	na
Physical well-being	0	na	na	na

na = not applicable/not studied

1. The number of centers represented in the communication/language competencies domain is an estimate. Crain-Thoreson and Dale (1999) do not report the number of classrooms assigned to each condition (three districts, five schools, and five classes were represented in the study). The design of the study makes it likely that all five classrooms were represented in each study condition.
2. A rating of “medium to large” requires at least two studies and two schools across studies in one domain and a total sample size across studies of at least 350 students or 14 classrooms. Otherwise, the rating is “small.” For more details on the extent of evidence categorization, see the WWC Procedures and Standards Handbook, Appendix G.